Journal of
Experimental Psychology

VOL. 26, NO. 6 JUNE, 1940

STUDIES OF ABNORMAL BEHAVIOR IN THE RAT. III. THE DEVELOPMENT OF BEHAVIOR FIXATIONS THROUGH FRUSTRATION *

BY NORMAN R. F. MAIER, NATHAN M. GLASER AND JAMES B. KLEE

University of Michigan

INTRODUCTION

Many investigators have observed that the pre-solution period of a discrimination problem is often accompanied by persistent position habits (*e.g.*, see Spence, **15**). Recently Maier and Glaser (**10**) and Maier (**12**) have pointed out that a discrimination problem if made insoluble results in position habits which are so strong that they can be abolished only with great difficulty. Since a problem which cannot be solved results in continued frustration, it is of importance to examine these position habits to determine whether they are the product of ordinary mechanisms involved in learning or whether they are an expression of mechanisms related to abnormalities in behavior. We speak of fixation in learning, meaning that habits have been well established, and we speak of fixations in abnormal psychology and mean that non-adaptive forms of behavior have become persistent parts of behavior. In the one sense fixation is a term used in normal psychology; in the other it refers to something abnormal.

It is the purpose of the present study to compare the strength of the same response when produced under the

* This study was supported by grants to the senior author from The Horace H. Rackham School of Graduate Studies of the University of Michigan and from the John and Mary R. Markle Foundation of New York City. The authors wish to express to the donors their sincere appreciation for this assistance.

ordinary conditions of learning and when produced under conditions of frustration and see whether the resulting fixations are of the same order.

METHODS

The Lashley jumping apparatus was used for establishing the position habits. The essential feature of the technique is to require the animal to leap from a jumping stand at cards placed in the two windows cut in an upright screen. These cards may be latched or unlatched. If the unlatched card is struck it falls over and the animal gains access to food placed on a platform behind the window, but if the latched card is struck the animal receives a bump on the nose and falls into a net below. The distance between the jumping stand and the card was 9 inches in these experiments.

Preliminary Training

To induce animals to jump at the cards in this apparatus requires some preliminary training. The procedure used in these experiments was adapted to suit the purposes of the experiments.

After having been fed on the food platform of the apparatus for the three preceding days, the first stage of training was to place the jumping stand close to the open windows of the apparatus so that the animal could walk across. Gradually the stand was moved back until the animal was required to leap across the gap. The number of trials required to cause the rat to make the leap varied from animal to animal, but in general from 2 to 3 days of 10 trials per day were adequate to accomplish this. Guidance with the hand was used to cause the animals to jump through both windows.

The second stage consisted of placing cards (15 cm. sq.) in the windows. One card contained a black circle (8 cm. in diameter) on a white background, the other a white circle (of the same size) on a black background. These were changed from side to side in an irregular order. During the first trials the cards were so placed as to cover only part of the window and gradually more and more of the card was exposed. Since both cards were unlatched the cards readily fell over whenever struck by the rat. Soon the window could be covered completely by the card and the rat responded by leaping at the card. To prevent the animal from jumping consistently to one side or to one of the cards manual guidance was introduced whenever the animal began to express a preference for one side or one card. This guidance consisted of pushing the rat away from the preferred window or card as soon as it prepared to jump and was continued until each rat readily reacted by jumping at the cards, which now completely covered the windows. In general 2 or 3 days of training were adequate to accomplish this. In a few cases the first stage was repeated if guidance failed to overcome position habits.

During the third stage the procedure consisted of guiding the animal's choices so that an equal number of responses was made to each card and to each position. The choice was thus determined by the experimenter. This procedure was continued for 4 days and 10 trials were given daily. In this stage all animals received identical treatment.

With the preliminary training completed the rats were divided into 3 groups and each was given special training in the establishment of a position habit. This special training was continued until 98 percent of the last 160 trials were responses to position. Thus the training extended over a period of 8 days or more and 20 trials were given per

day. Preliminary experimentation demonstrated that approximately 160 repetitions of the position habit were adequate for producing group differences.

With the introduction of special training, air led by a tube to the jumping stand was used to induce the animals to jump. The air was applied according to the following schedule: (1) no air for 30 seconds, (2) mild air blast for 30 seconds, (3) medium air blast for 30 seconds, and (4) full air pressure (10 to 14 pounds) for the remainder of the period of resistance. (The intensity of the air was controlled by marked positions on the valve.) As soon as the animal made a choice the air was turned off and the rat's resistance was rated according to the amount of air required to induce a response.

Special Training

Group I. Position Habit—Preferred Side.—The rats in this group, after choosing either the right or left card on their first free choice after the above preliminary training, were forced to continue to respond to the same side. This was done by latching the window on the side opposite to that chosen on the animal's first jump (*e.g.*, if an animal jumped through the right window on its first jump to a card the left window was thereafter latched). Thus a jump to the other side became associated with striking a card fixed in position with the resulting bump on the nose and fall to the net below. The cards were changed from side to side in an irregular sequence so that position rather than the character of the card became associated with success.

Group II. Position Habit—Reversed Preference.—This group of rats was treated in a manner similar to Group I except that both windows were latched on the first jump and the side to which the rat jumped became the latched side thereafter. If the first jump is considered to be an expression of a natural preference the rats in this group were required to form a position habit which was opposed to their natural preferences, whereas the rats in Group I formed position habits which coincided with their natural preferences. In both groups a consistent response to position was rewarded 100 percent of the time.

Group III. No-solution Problem.—In this group the experimenter locked one of the windows at the outset. Thereafter either the right or left window was latched in a random order. As above, the cards were changed from side to side in an irregular sequence. In each series of trials the two sides and the two cards were latched and unlatched an equal number of times. This group was thus rewarded on half its trials and punished on the other half no matter on what basis the choices were made. This situation readily resulted in the appearance of position habits in 11 out of 13 animals. The two others reacted on the basis of the characteristics of the card and were excluded to make possible comparisons with Groups I and II.

Tests on the Persistence of the Position Habit

Since each of the groups developed position habits the following tests were designed to measure the strength of the habit in each of the groups.

Discrimination Problem.—Upon completion of the special training all rats were presented with a discrimination problem. The card with the black circle on the white background was made the positive stimulus and a response to it led to reward whether on the right or left side, whereas the card with the white circle on the black background was made negative and a response to it always led to punishment. Under ordinary conditions of learning the average rat in our colony requires 88 trials before making

30 consecutive errorless jumps. To allow ample time for learning the three groups of rats were therefore given 200 trials (20 trials per day) to learn this problem. If the animals persisted in their position habits under these conditions they received punishment on 50 percent of their trials. For Groups I and II this would result in a change in their pattern of punishment and for Group III the proportion of punishment would remain the same but punishment would now occur in an orderly fashion.

Position Response Consistently Punished.—Animals failing to show evidence of learning the discrimination problem were transferred to a new problem in which the negative card was always on the side to which they reacted consistently and of course this card was always latched (*e.g.*, an animal with a right position habit always received punishment when it chose the negative card on the right). This problem was continued for 10 days and 10 trials were given daily. Clearly this manner of punishment should have discouraged the position habit and encouraged the formation of a response based upon the characteristics of the card.

Subjects Used

Complete records were obtained on a total of 31 rats. Groups I and II each contained 10 animals and Group III contained 11. The age range was from 4–9 months and the average age for Groups I, II, and III was 5.4 months, 6.4 months, and 5.6 months respectively. Of the 31 animals 21 were males and 10 were females. The proportions of males to females for Groups I to III were 7 : 3, 6 : 4, and 8 : 3, respectively. Both white and pigmented animals chosen from the regular laboratory colony were used.

Results

Comparison of Groups in the Persistence of the Position Habit

Maier (11) has pointed out that the fixation of a response may be measured by the resistance which a given mode of behavior shows to modification. Since all of our groups of rats developed position habits and since this mode of response was repeated without alteration for about 160 consecutive trials (the criterion was a response to a position in 98 percent of the last 160 trials) these groups may be compared in their degree of fixation. Because the introduction of the discrimination problem makes possible a new form of behavior which is adaptive in nature one might expect the adaptive response to be learned and the unadaptive response to be replaced by it. If such a change in the habit fails to take place one might expect that 100 percent punishment of a fixated response would cause it to be abandoned. If both treatments fail to cause a given mode of behavior to be abandoned one must recognize the existence of a kind of fixation which requires careful analysis since it is not deducible from our knowledge of learning.

Table 1 shows that all but one of the rats in Group I abandoned the position habit during the 200 trials but only 4 of the 11 rats in Group III made this adjustment. The rats in Group II fall between these extremes but are more similar to Group III than to Group I.

The females seem to have a greater tendency than males to persist in their position habits since 6 out of 10 females persisted and only 7 out of 21 males did so. Since Group III had the smallest percentage of females, the difference between groups II and III is perhaps greater than shown in Table 1. Considering males only, 1 of 7 in Group I, 2 of 6 in Group II and 4 of 8 in Group III persisted.

If the animals in each of the groups differ in the relative amount of persistence, then more drastic treatment should

TABLE 1

COMPARISON OF GROUPS IN EXTENT OF FIXATION

Group	Number of Rats Learning Discrimination	Number of Rats Failing to Learn Discrimination	Number of Rats Persisting after 100 Percent Punishment
I	9	1	1
II	5	5	5
III	4	7	7

cause additional animals to abandon their position habits. The last column of Table 1 shows that this was not the case, however. Animals which failed to learn the discrimination continued to show a position response throughout the 100 punishment trials as well as during the 200 trials on the discrimination problem. This suggests an all or nothing character of the persistence of the positional response.

We may also study the animals which learned the problem and compare the groups with regard to the readiness with which these animals formed the discrimination habit.

From Table 2 we see that the position habit was abandoned after 55 trials in the case of Group I, after 49.8 trials in the case of Group II, and after 98.0 trials in the case of Group III. With respect to the animals which learned, Groups I and II are approximately alike and differ from Group III to the extent that the score ranges show practically no overlapping

TABLE 2

TRIALS REQUIRED TO LEARN THE DISCRIMINATION HABIT

Group	Percent Solving Discrimination Problem	Average Number of Trials in Which Position Habit Persists	Range	Average Number of Additional Trials to Reach Criterion (Exclusive of Criterion Trials)
I	90.0%	55.0	20–105	27.2
II	50.0%	49.8	21–80	16.2
III	36.4%	98.0	85–117	4.5

(only one case in Group I falls within the score range of Group III). In all cases, however, in which learning occurs it takes place well before the end of the 200 trials allowed for learning. Thus the animals which learn and those which persist in the position habit seem to differ qualitatively, a point which further substantiates the conclusion that persistence may take on an all or nothing character.

The groups also differ in the manner in which the new discrimination is established once the position habit is abandoned. This is shown in the last column of Table 2. Group I requires an average of 27.2 trials before it consistently reacts to the positive card; Group II requires 16.2 trials; and Group III requires only 4.5 trials. Thus Group III shows sudden learning after abandoning the position habit, despite the fact that it is more reluctant to abandon it.

To throw further light on the group differences as well as on the question of the all or nothing character of the learning within the groups let us turn to the original learning of the position habit. These data are shown in Table 3. It will

TABLE 3

TRIALS REQUIRED TO FORM POSITION HABIT

	Average Number of Trials Before Criterion is Reached (98 Percent in 160 Trials)		
	All Animals	Animals Solving Discrimination Problem	Animals not Solving Discrimination Problem
Group I	0	0	0
Group II	9	8	10
Group III	20	7.5	27.1

readily be seen that Group I, whose position habit corresponded to its preference, continued to react on the basis of this preference without ever adopting another mode of response. Group II on the average required 9 trials before it settled down to a positional response opposite to its first reaction. Group III, which was punished on half its trials no matter upon what its reaction was based, reacted in terms of position after 20 trials.

We may therefore say that Group III had the greatest opportunity to eliminate alternative modes of behavior (hypotheses) before consistently reacting to position. The position habit thus becomes a last resort and resists being displaced by another reaction tendency. The elimination of certain behavior tendencies at the outset explains why Group III rats are the slowest learners on the discrimination problem and may also explain why this group has the smallest number of learners. The fact that the animals of this group which solved the discrimination problem formed their position habit after 7.5 trials and those which failed to solve formed their position habit after 27.1 trials supports this interpretation. The explanation for the sudden learning of the animals which learned the discrimination in Group III is also simplified if we grant that certain behavior tendencies have already been eliminated. Krechevsky (**9**) has shown that during the pre-solution period a problem may be changed without reducing the final rate of learning.

A similar analysis may be applied to Group II. The average number of trials preceding the formation of the position habit for the animals in this group lies between that of Groups I and III and their tendency to persist in the position habit also lies between these two groups. The learners and non-learners, however, are not differentiated here.

Evidence for Suppressed Learning of the Discrimination

As already pointed out the introduction of the discrimination problem changed the pattern of punishment and reward for Groups I and II but since Group III was previously punished half the time the proportion of reward and punishment

remained the same for it. Thus it might be said that Group III tended not to learn the discrimination problem because the opportunity for learning did not become apparent and hence there was no reason for abandoning the mode of behavior which was established. If this explanation is sound then Groups I and II should have learned the discrimination problem equally well. Further Group III should have abandoned the position habit when 100 percent punishment of the position habit was introduced since this condition changed the pattern of punishment. On both of these counts the experimental findings are in disagreement with the demands of this theory.

As a matter of fact it can be demonstrated that the animals which failed to abandon their position habits actually learned the discrimination.

Throughout the experiment the resistance to jumping was overcome by the application of the air blast. By increasing the air intensity at certain intervals, ratings for resistance could be made on the basis of the length of resistance and air intensity used before the jump occurred. The following ratings for resistance were used:

Rating 0.	No air.	Jump occurring between 0– 30 sec.
Rating 1.	Mild air blast.	Jump occurring between 31– 35 sec.
Rating 2.	Mild air blast.	Jump occurring between 36– 60 sec.
Rating 3.	Medium blast.	Jump occurring between 61– 90 sec.
Rating 4.	Full blast.	Jump occurring between 91–120 sec.
Rating 5.	Full blast.	Jump occurring between 120–150 sec.

If a rat always jumped to one side but consistently showed greater resistance to jumping when one of the cards was on the position-response side than when the other card was on that side, then a discrimination between the cards must have been made. We have therefore analyzed the resistance shown to the positive and negative cards during the discrimination test to determine whether the animals which failed to choose the positive card expressed greater resistance for the negative than for the positive card.

For comparison we have also studied the resistance to the two cards during the special training period to see whether a preference for one of the cards was evident before reward and

punishment were introduced in connection with one of the cards. These results are shown in Table 4. For each group of rats the average resistance to each card as well as the difference in resistance is shown. It will be seen that all 3 groups failed to show differential resistance to the two cards during the special training period in which the position habit was

TABLE 4

RESISTANCE BEHAVIOR OF RATS PERSISTING IN POSITION HABITS

Number of Trials	Group I (1 rat) Average Resistance Rating			Group II (5 rats) Average Resistance Rating			Group III (7 rats) Average Resistance Rating		
	Positive Card	Negative Card	Difference	Positive Card	Negative Card	Difference	Positive Card	Negative Card	Difference
Special Training Period									
160	0	.01	.01	.57	.55	−.02	.82	.91	.09
Discrimination Problem. Position Habit Punished on Half the Trials									
1–40	2.00	2.05	.05	1.05	1.07	.02	.89	.95	.06
41–80	1.60	3.10	1.50	.76	1.41	.65	.68	1.04	.36
81–120	1.55	3.15	1.60	.97	1.72	.75	.65	1.52	.87
121–160	.55	3.80	3.25	1.17	1.74	.57	.80	1.66	.86
161–200	.40	3.10	2.70	.80	1.88	1.08	.68	1.62	.94
Average	1.22	3.04	1.82	.95	1.56	.61	.74	1.36	.62
Position Habit Punished on All Trials									
1–40	Data not available			1.72			1.58		
41–80				1.78			1.82		
81–100				1.42			2.36		
Average				1.68			1.83		

being established. The groups differ somewhat in the amount of resistance, however. Group I had a rating of practically zero which means that this group responded before air was applied. However, the data on this group have limited value since they are based on one rat, all the others having learned the problem. Group II with a rating of a little over .5 required air about half the time and Group III, which was punished on half its trials, had a resistance rating of nearly 1

which means that on the average these rats jumped as soon as air was applied.

As soon as the discrimination problem was introduced all groups of rats were punished on half the trials until they solved the discrimination problem. As a consequence resistance immediately rose for rats in Groups I and II which experienced a change in the punishment pattern and remained about the same for Group III which had no change in the frequency of punishment.

With the introduction of the discrimination problem the black circle on the white card became the positive stimulus and the white circle on the black card became the negative stimulus. Since all rats consistently jumped to the right or left side they were punished whenever the negative card appeared on the side of the positional response. In order to show the change in the relative resistance to jumping at the positive and negative cards the data on resistance are given for each 40 trials.

From Table 4 it will be seen that during the first 40 trials the resistance to the two cards was about the same. The next four groups of 40 trials, however, show a change in the relative resistance which is maintained throughout the remaining 160 trials. The difference in reaction in all groups is due in part to the increase in resistance for the negative card and in part to a reduction in resistance for the positive card.

When the position habit is punished 100 percent of the time rather than 50 percent, as was the case during the discrimination problem period, only a slight increase in resistance to jumping to the negative card occurs despite the fact that the expectancy of punishment for each positional response increases. Brunswik (1) has shown that the probability of reward and punishment is a determining factor in learning and under ordinary circumstances such a change in procedure should have an influence on behavior.

The data on resistance analyzed above show that a discrimination response has been superimposed upon the position habit. Apparently a choice based on the difference between the cards is not able to replace the position habit and thus the

differentiation expresses itself in the resistance to jumping to the negative card.

For rats solving the problem the data on resistance must be treated somewhat differently since the number of trials to solve varied from individual to individual. We have, therefore, broken the discrimination period into three parts: (*a*) the pre-solution phase during which the position habit persisted; (*b*) the solution phase during which the position habit was abandoned and before the positive card was consistently chosen; and (*c*) the criterion phase during which the positive

TABLE 5

RESISTANCE BEHAVIOR OF RATS WHICH ABANDONED THE POSITION HABIT FOR THE DISCRIMINATION HABIT

Number of Trials	Group I (9 rats) Average Resistance Rating			Group II (5 rats) Average Resistance Rating			Group III (4 rats) Average Resistance Rating		
	Positive Card	Negative Card	Difference	Positive Card	Negative Card	Difference	Positive Card	Negative Card	Difference
Special Training Period									
160.....	.30	.31	.01	.30	.45	.15	.88	.92	.04
Discrimination Problem. Position Habit Punished on Half the Trials									
Pre-solution phase	.88	1.05	.17	.63	.99	.36	.89	1.34	.45
Solution phase....	1.00	.60	−.40	.93	1.25	.32	.88	0	−.88
Criterion phase...	.67			.65			1.13		

card was consistently chosen. Since these animals abandoned the position habit and learned the discrimination they were not presented with the period in which the position habit was punished 100 percent of the time.

The data for these animals are given in Table 5. As in the case of rats not solving the discrimination problem the resistance behavior toward the two cards is about the same during the special training and is higher for rats in Group III which received punishment for their position responses.

With the introduction of the discrimination problem only a slight difference in resistance for the positive and negative cards appears during the period in which the position habit

continues to function. Table 5 shows these differences to be .17, .36, and .45, for Groups I, II and III respectively, whereas Table 4 showed differences of 1.82, .61 and .62 for the non-solving animals in these same groups.

During the solution period the resistance to the negative card rises for Group II but falls in the cases of Groups I and III. However, these differences in resistance to the two cards have little significance since very few responses were made to the negative card after the positional responses were discontinued. As shown in Table 2, Group I, after abandoning the position habit, learned in an average of 27.2 trials; Group II in 16.2 trials; and Group III in 4.5 trials. As a consequence of this rapid learning the average number of responses to the negative card during the solution period was 3.6 for Group I, 1.6 for Group II, and .25 for Group III. Thus we see that although the discrimination learning was slightly expressed during the pre-solution phase the position habit was discontinued before the differential resistance became as pronounced as in the case of animals failing to solve. This fact clearly shows that the animals which failed to solve the discrimination problem failed because they were unable to abandon their position habits and not because they did not sufficiently differentiate between the positive and negative cards.

After the animals solved the problem and consistently chose the positive card (the criterion phase) the resistance to jumping showed no change of sufficient magnitude to require elucidation.

Evidence for discrimination on the part of animals failing to abandon the position habit may also be derived from their manner of jumping to the negative and positive cards. Abortive jumps, consisting of (*a*) jumps to the ledge, (*b*) striking the card with the side of the body, (*c*) leaps to the right or left of the cards, and (*d*) jumps which are very light and not adequate for knocking over an unlatched card, often replace the normal response which consists of a fairly vigorous leap at the center of the card. We have therefore analyzed the data to determine the percentage of abortive responses to each of the cards. These data are presented in Table 6.

This table shows that during the special training abortive behavior was practically absent with the exception of Group III, which received punishment on half the trials and in which nearly 10 percent of the responses were abortive. With the introduction of the discrimination problem abortive behavior became more common and became applied to the negative

TABLE 6

ABORTIVE BEHAVIOR OF RATS PERSISTING IN POSITION HABITS

Number of Trials	Group I (1 rat) Percent Abortive Jumps			Group II (5 rats) Percent Abortive Jumps			Group III (7 rats) Percent Abortive Jumps		
	Positive Card	Negative Card	Difference	Positive Card	Negative Card	Difference	Positive Card	Negative Card	Difference
Special Training Period									
160	0	0	0	4.8	4.1	−.7	9.7	9.6	−.1
Discrimination Problem. Position Habit Punished on Half the Trials									
1–40	0	0	0	19	13	−6	13	14	1
41–80	0	0	0	5	4	−1	0	11	11
81–120	0	20	20	7	27	20	0	19	19
121–160	0	95	95	1	58	57	0	23	23
161–200	0	95	95	1	54	53	1	37	36
Average	0	42	42	6.6	31.2	24.6	2.8	20.8	18.0
Position Habit Punished on All Trials									
1–40	Data not available			71			32		
41–80				90			41		
81–100				96			42		
Average				83.6			37.6		

card. Groups I and II showed a differential response to the two cards after 80 trials and Group III after 40 trials. Thus all groups expressed a discrimination between the two cards by developing an abortive response for the negative card although this expression of the discrimination occurred somewhat later than the differential resistance (see Table 4).

With the introduction of the period of 100 percent punishment Groups II and III showed an increase in abortive behavior and for Group II it rose as high as 83.6 percent.

These data are not available for Group I, and since only one rat in this group failed to solve the problem the figures for Group I should be taken as an individual rather than as a group trend.

It appears that abortive behavior is a form of adjustment to the situation. The fact that Group II showed a high percentage of such behavior in the 100 percent punishment series with very little rise in resistance, whereas Group III showed a lesser percent of abortive behavior and a significant rise in resistance during this period, supports this contention.

TABLE 7

ABORTIVE BEHAVIOR OF RATS WHICH ABANDONED THE POSITION HABIT FOR THE DISCRIMINATION HABIT

Number of Trials	Group I (9 rats) Percent Abortive Jumps			Group II (5 rats) Percent Abortive Jumps			Group III (4 rats) Percent Abortive Jumps		
	Positive Card	Negative Card	Difference	Positive Card	Negative Card	Difference	Positive Card	Negative Card	Difference
Special Training Period									
160.....	.3	0	−.3	1.5	1.5	0	2.2	2.9	.7
Discrimination Problem. Position Habit Punished on Half the Trials									
Pre-solution phase	4	5	1	1	0	−1	23	28	5
Solution phase....	1	0	−1	0	0	0	0	0	0
Criterion phase...	3			0			0		

The abortive behavior of animals solving the discrimination problem is shown in Table 7. As in the case of animals failing to solve this problem these animals show practically no abortive behavior during the special training but unlike them, these animals do not develop it as a differential reaction when the discrimination problem is introduced. Groups I and II apparently learn the discrimination before abortive behavior appears and Group III (which required about twice as many trials as Groups I and II for learning the discrimination (see Table 2)) shows abortive behavior for both cards but this behavior drops out as evidence for learning appears. Apparently the abortive behavior appears only after a certain

amount of frustration and ceases as soon as an adjustment occurs. The animals which fail to solve persist in abortive behavior not because they cannot react differentially to the two cards but because they cannot achieve an adjustment as long as the position habit dominates their behavior.

The evidence on resistance and abortive jumping shows that the animals failing to solve the discrimination problem actually formed the discrimination and expressed the discrimination by reacting differently to the positive and negative cards. Of the 13 animals in the three groups which failed to abandon the position habit, 9 expressed their differentiation both in their resistance and in their abortive behavior, 3 in their resistance behavior only, and 1 in its abortive behavior only. In no case did an individual fail to express a differentiation between the cards.

Although a variety of forms of abortive behavior appeared in these experiments each animal developed its own individual abortive pattern to which it strictly adhered.

Additional Experiments on Frustration

Thus far our experiment has been carried to the point of demonstrating the unusual persistence of a habit produced by frustration. To throw further light on behavior arising from frustration, 4 additional animals were studied over a longer period of time. During this period the effects of breaking the position habits was also studied.

In 2 animals (rats 1 and 2) the no-solution problem was introduced after the animal was trained to jump and was continued for 450 trials (20 trials per day). One of the cards contained a large black circle and the other a small black circle and both were on gray backgrounds. The position of the cards was changed irregularly and either one or the other was latched on each trial. One rat immediately took a left position habit going left 9 times in the first 10 trials and to the left on the remaining 440 trials. The other likewise chose the left 9 times on the first 10 trials, then chose both sides, but after 30 trials it always chose the card to the right.

After this period of frustration the card with the large

circle was made positive. For the next 460 trials both rats continued to choose on the basis of position, never once deviating from this mode of behavior.

The window opposite the position preference was then left without a card and the rats were induced by means of guidance to jump through the open window. After 20 such trials the positive card (large circle) was placed in the open window. Both rats chose the positive card and continued to choose it during the next 120 trials regardless of the side on which it was placed.

The no-solution problem was then introduced again and for 200 trials both of the rats consistently chose the positive card of the discrimination.

In this experiment persistence behavior was shown to appear in the no-solution problem on both occasions, but on the first occasion the position habit persisted and on the second the discrimination persisted. Thus persistence may be shown for an obviously acquired preference as well as for what may appear to be a natural preference, and hence the persistence behavior must be regarded as a product of frustration rather than the expression of any natural strength the position habit may possess. In the previous experiment it was shown that there was greater persistence for a position habit which opposed the animal's natural preference than for one which corresponded to it.

To further isolate the causes that determine which habit will persist the remaining 2 rats (numbers 3 and 4) were tested in the following manner.

A discrimination between the small black circle and the large black circle on gray backgrounds was established, the former being the positive stimulus. After the discrimination was established to the criterion of 30 errorless trials the no-solution problem was introduced. Rat 3 lost the discrimination habit after 140 trials and Rat 4 after 40 trials and both reacted thereafter by showing left positional reactions. Rat 3 was given 250 trials and Rat 4, 80 trials, and both consistently reacted on the basis of position during this period.

The original discrimination problem was then reintroduced.

Rat 3 continued its position habit for the 330 trials it was tested and Rat 4, for the 150 trials it was given.

Since neither of the animals deviated from the position habit, the response was broken by means of the open window procedure described above. As soon as the positional response was abandoned reactions on the basis of the discrimination appeared without any pre-solution trials. Rat 3 was given 80 repetitions and Rat 4, 150, and both consistently chose the card with the small circle.

The no-solution problem was again introduced. Both rats failed to abandon the now irrelevant discrimination habit although Rat 3 was given 450 trials and Rat 4, 630 trials.

The experiments on Rats 3 and 4 show that frustration produces a position habit even when preceded by a discrimination. It is only when the fixation of the position habit has been broken that the rats fail to return to it. In such cases frustration results in the persistence of some other mode of behavior. In both forms of behavior the persistence is abnormally strong and can be broken only by special means. Ordinary conditions for learning do not furnish the necessary instigation for abandoning the habit which becomes fixated. It is only when an alternative is specifically presented by the experimenter (guidance) that a shift in response is made. The animals' own resources apparently are not adequate for taking an alternative even when the inadequacy of a response is experienced repeatedly. This conclusion is supported by observations on additional rats not included in this study.

Discussion

The Nature of Abnormal Fixations

We have found that some animals show a marked reluctance to abandon their position habits. This persistence in a mode of behavior has an abnormal appearance for the following reasons:

1. Animals persisting and those not persisting form a distinctly bimodal distribution in the sense that all of the latter abandon the position habit in 117 trials or less and all of the former fail to abandon it in 200 trials. Even the addition of

100 trials of 100 percent punishment fails to destroy the persistence of the position habit.

2. Inability to distinguish between the positive and negative cards does not account for the failure to learn the discrimination habit. The analysis of resistance and of abortive behavior revealed that each animal which failed to form the discrimination habit actually learned the discrimination. Failure to express this discrimination seemed to be due to the dominance of the position habit. Thus the discrimination, even when present, was unable to break through the position habit.

3. Ordinarily the strength of a habit varies with the frequency with which a response has been rewarded. Brunswik (1) has demonstrated varying rates of learning by varying the proportion of reward for the positive card and punishment for the negative. In the present experiment the condition of 50 percent reward and 50 percent punishment for the position habit produced a position habit of greater tenacity than did an equal number of trials with 100 percent reward. Such a condition should produce no learning if we are to accept the principles of learning supported by many studies. Hence the fixation here produced cannot be of the same kind as the fixation produced under ordinary conditions of learning.

To reconcile these results with those of other experiments it is necessary to postulate a different mechanism which has been brought to expression. Our Group III and to some extent our Group II animals were in a frustrating situation. If we suppose that frustration may lead to fixation, this fixation may be of a qualitatively different nature than that produced by learning. In the sense that such fixations are not the ordinary ones produced by learning they may be called abnormal fixations. Frustrations also lead to other forms of behavior which have been called abnormal (and in a sense extreme resistance and abortive jumping are abnormal) and it therefore seems justified to so classify these fixations. Case histories of neurotic people contain much evidence of persistence of unsuccessful modes of behavior and when these arise through frustration they are readily classified as belonging to the abnormal.

4. Although positional responses usually occur in frustrating situations, fixations are not limited to this mode of behavior. The experiment on the last four rats shows that a discrimination habit may also become fixated but this occurs only after the position fixation has been abandoned. Once a fixation is abandoned there seems to be no tendency to return to it. Rather the fixation shifts to another mode of behavior which is in the animal's repertoire. In each case the basic cause of the abnormal fixation seems to be the no-solution problem which is frustrating.

5. The abandonment of a fixation occurs suddenly if at all and then it constitutes a shift to a different mode of behaving. This fact is illustrated by the behavior of the four animals which were extensively studied and by the sudden learning of the Group III animals which abandoned their position habits. The corresponding animals in Groups I and II which were frustrated to a lesser degree abandoned their position habits after experiencing failure and before a new habit was found. They therefore required trials in which to learn a new response. Thus the animal which forms an abnormal fixation retains its mode of behavior until another is substituted. In this sense the fixation is an adjustment which the animal retains and as a consequence it is incapable of further learning by trial and error under the usual conditions for learning.

6. Because all pre-solution periods in learning are in essence no-solution problems for a time one might expect fixations to occur in all learning situations. This, however, is not the case. Apparently frustration leads to fixation only when it reaches a certain point (let us say a point of saturation for that individual). To put it differently, continued failure causes the animal to 'give up' and when this point is reached the mastery of a new problem no longer takes place. Instead the animal has formed an adjustment which is a fixed mode of behaving.

That individual differences occur is apparent from our results. All animals in Group III did not develop fixations and one animal in Group I did form a fixation. Everall (3)

pointed out that some animals perseverate no matter what the conditions are and Hamilton and Krechevsky (**6**) selected their animals to avoid such individuals. Because some animals are so prone to fixation we feel justified in confining our discussion to the role played by the situation in producing fixations.

Since we believe that frustration must reach a saturation point in order to become effective, we are inclined to disagree with the stimulating volume by Dollard, Doob, Miller, Mowrer, and Sears (**2**) in so far as these authors insist that all unrewarded experiences are frustrating. We prefer to limit the term frustration to designate the state where continued failure causes learning functions to cease operating and other mechanisms of adjustment to begin operating. In this sense abnormal behavior becomes qualitatively different from normal.

Fixation and Regression

Although our experiments have been confined to the study of fixations, they show a similarity to some recent experiments on regression in which punishment caused rats to abandon one mode of behavior and fixate a previously learned habit. The strength of the fixation was not determined in these studies, but rather they were pointed to bring out the regressive feature of behavior. Since our experiments demonstrate the continuance of a given mode of behavior during frustration, it is desirable to compare methods and results to see if the apparent difference in the findings can be clarified.

Hamilton and Krechevsky (**6**) developed in rats a preference for a right turning response in a simple maze and then altered the situation and developed a left turning response. When shocked at the choice-point 11 out of 18 rats consistently chose the right turn as in the first habit and discontinued the second habit. When animals were shocked with no previous training a good proportion of them showed consistent position preferences. These authors speak of the fixation of the earlier habit as regression. Everall (**3**) raised the question whether the reversion was to an earlier response or to some natural preference and concluded the former to be the case.

Sanders (**14**) used a similar procedure and in addition to verifying the earlier work found that the relative strengths of the two position habits had no effect on the tendency to revert to an earlier habit. Furthermore shocking the rats during the second habit caused reversion to the first but when the first habit was again established the introduction of electric shock failed to cause a reversion to the second habit. She therefore believes that regression rather than disinhibition is demonstrated by these experiments.

The fixation of behavior described in the above experiments has been produced by electric shock and the authors believe the emotional effect produced by the shock to be responsible. Hamilton (**5**), some years ago, emphasized the perseverative effect of emotional excitement on behavior. Recently Patrick (**13**) verified Hamilton's work by demonstrating that college students become less rational and more random and stereotyped during emotional excitement. Our experiments have likewise demonstrated fixation but the causal factor seems to be frustration. It is possible that emotion is involved in both cases but in our experiment the accumulative effect rather than the direct arousal of intense emotional excitement has been the dominant feature. Furthermore our fixations have shown greater resistance to change. Even when animals developed a discrimination and had an adaptive mode of behavior available the fixated response dominated the choices of the animals.

It is not our opinion that we have demonstrated regression in behavior. Rather we have found that frustration causes a position habit to become fixated. Once a position fixation is broken, however, another mode of behavior can become equally fixated and in this respect our findings differ from those of Sanders. Any situation which disrupts one mode of behavior must give rise to another and this other available mode of behavior may be a previous one. The appearance of a previous mode of behavior, however, is not necessarily evidence for regression if by regression one wishes the term to imply a reversion to a more primitive behavior tendency. To demonstrate that the reversion is a return to primitive be-

havior supplementary evidence should be presented which indicates that the behavior is not only an earlier form but a more primitive form.

Because the work on regression has dealt with position habits and because we have found that these become formed and fixated in a frustrating situation, the appearance of a position habit may not be a return to an earlier response but rather the formation of a new one which happens to correspond with the position habit previously formed.

In all of the above experiments animals learned a right position habit and then a left position habit. When shocked they responded with the right position habit. If another group had formed the left position habit only and if it was then shocked, would this group also respond by going to the right in the same proportion? If so the shock would seem to cause the abandoning of one position habit and the appearance of a new one. The appearance of the right position habit in the first case would then not seem to be a return to an earlier habit but rather the formation of a new habit which happened to correspond with the first.

Sanders' inability to cause a return to the second habit may be due to the fact that the same animals had previously abandoned this habit in the punishment situation. This could be tested by developing a left position habit, then, a right, and then retraining on the left habit. After this retraining, would shock cause a right going habit to appear? According to Sanders' interpretation, the left habit should continue since the left habit was the earliest. We suspect that in this case the right habit would appear.

Of course one may say that all animals have natural position preferences which are the earliest and most primitive modes of behaving. Hence the formation of a position habit is a regression to these early tendencies. We have shown, however, that the size discrimination habit may become fixated and this form of reaction is far from primitive since such a habit is learned by rats with considerable difficulty. We therefore believe that fixation but not necessarily regression has occurred in our experiments, and it is possible that the

same interpretation holds for the punishment experiments since the regressive feature may be an artifact due to the use of position habits. In any case the relationship between fixation and regression must be studied further and for the present we wish only to make a case for fixation.

Humphreys (7) has recently shown that a conditioned eyelid response in the human is more difficult to extinguish if it is formed on the basis of 50 percent reinforcement than if formed on the basis of 100 percent reinforcement. This is true even when the number of reinforcement trials are equated in the two conditions. Humphreys (8) explains this on the basis of expectancy of reinforcement. A response which is reinforced part of the time continues to be effective when reinforcement is completely withdrawn because reinforcement is not always expected and hence failure is not demonstrated. But when reinforcement is withdrawn from individuals which have experienced reinforcement on all occasions a change in experimental conditions is more readily apparent to the subject. The mechanism of expectancy could explain why our Group III animals which did not fixate learned the discrimination less readily than the corresponding animals in Groups I and II but it would not account for our fixations.

Humphreys' findings are consistent with the implication that persistence is due to the tendency not to vary behavior. Maier (11) has pointed out that behavior may continue unaltered either because variability in behavior is not called out in the situation or because a particular response tends to dominate over other modes of behavior. The latter form of behavior should be called stereotypy since it dominates because of its own properties rather than because other behavior alternatives were not called out. Humphreys' form of persistence is evidence for a response which continues because nothing in the situation calls out a different form of behavior. In our experiments we have shown that another form of behavior (discrimination) had actually been aroused but failed to come to complete expression because the position habit could not be replaced. As a consequence the new form of behavior became superimposed upon the position habit. We

are therefore inclined to believe that Humphreys was not dealing with an abnormal form of fixation. This is also shown by the fact that all his subjects finally did extinguish the conditioned response, and only additional trials were required by the 50 percent reinforcement group.

Gottsdanker (4) studied the fixation of college students in a multiple choice problem in which each of a certain group of keys was correct 75 percent of the time. He found that blindfolded students were more disposed than non-blindfolded ones to confine their choices to a particular key rather than to the group of keys which had the same probability of being correct. He concluded that the fixation was 'explainable on the basis of lack of differentiation among choice possibilities on the part of the subjects.' Since our animals which fixated did show a differentiation between the choice possibilities this explanation does not hold for our data.

Because fixation is both a concept in learning and in abnormal behavior and because the mechanisms upon which these fixations depend are very probably quite different, it is important to differentiate between them. The belief that fixations appearing in certain experimental conditions invalidate the laws of learning is reasonable only when such a distinction is not made. The postulation of two kinds of fixations seems more sound than the throwing aside of a large body of concepts developed through ordinary learning experiments.

Summary

In this experiment rats were forced, by means of an air blast, to choose between 2 cards in the Lashley jumping apparatus. When the problem was insolvable position habits were formed, the strengths of which were greater than those of position habits formed under the usual conditions of learning.

A study of the persistence of position habits formed under 3 experimental conditions revealed that animals in each group could be classified into (*a*) those which solved a new discrimination problem readily and (*b*) those which failed to solve the new problem in more than ample learning time.

Even the addition of 100 trials of 100 percent punishment failed to cause a displacement of the persistent type of position habit. Since no intermediate cases were found it became evident that we were dealing with a discontinuous distribution of individual characteristics. This led us to postulate that the fixations in many individuals take on an abnormal character.

The data were then analyzed to determine the cause of the persistence of the position habits. It was shown that the animals which fixated actually solved the discrimination problem but nevertheless continued their positional behavior. In such cases the discrimination was evidenced by (*a*) the differential resistance to jumping to the positive and negative stimulus cards (when they appeared on the position side) and (*b*) the different manner (normal or abortive) in which the animals jumped at these two cards. Thus the discrimination habit became superimposed on a fixated position habit. This demonstrated that the persistence of the position habits was not due to the absence of alternative modes of behavior, but rather to the strength of the position habit.

Since all learning situations involve some frustration and since a no-solution situation involves greater frustration than solvable problems, the fixation behavior became explainable only when we postulated that fixation is the result of frustration. The individual differences were explained by considering frustration as accumulative. Only when failure is repeated to the point of the individual's level of tolerance does it constitute frustration; but once the level of frustration is achieved, fixation appears.

Under ordinary conditions the fixation expressed itself as a persistent position habit. It is only when the position fixation was broken that another mode of behavior became fixated. When this occurred the other fixation also was of abnormal strength.

The data were also considered in relation to regression produced in situations of emotional stress and although certain doubts concerning regression were entertained no definite conclusion was expressed.

The consequences of frustration must be distinguished from those of conflict since frustration as such does not appear to result in experimental neurosis.

(Manuscript received December 14, 1939)

References

1. Brunswik, E., Probability as a determiner of rat behavior, *J. exper. Psychol.*, 1939, **25**, 175–197.
2. Dollard, J., Doob, L. W. Miller, N. E. Mowrer, O. H., and Sears, R. R., *Frustration and aggression*, New Haven; 1939, Yale Univ. Press, pp. 209.
3. Everall, E. E., Perseveration in the rat, *J. comp. Psychol.*, 1935, **19**, 343–369.
4. Gottsdanker, R. M., An experimental study of fixation of response by college students in a multiple-choice situation, *J. exper. Psychol.*, 1939, **25**, 431–444.
5. Hamilton, G. V., A study of perseverance reactions in primates and rodents, *Behav. Monogr.*, 1916, **3**, pp. 65.
6. Hamilton, J. A., and Krechevsky, I., Studies in the effect of shock upon behavior plasticity in the rat, *J. comp. Psychol.*, 1933, **16**, 237–253.
7. Humphreys, L. G., The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions, *J. exper. Psychol.*, 1939, **25**, 141–158.
8. Humphreys, L. G., Acquisition and extinction of verbal expectations in a situation analogous to conditioning, *J. exper. Psychol.*, 1939, **25**, 294–301.
9. Krechevsky, I., A study of the continuity of the problem-solving process, *Psychol. Rev.*, 1938, **45**, 107–133.
10. Maier, N. R. F., and Glaser, N. M., Experimentally produced neurotic behavior in the rat, Film, Bethlehem: 1938, A. Ford, 600 ft 16 mm.
11. Maier, N. R. F., The specific processes constituting the learning function, *Psychol. Rev.*, 1939, **46**, 241–252.
12. Maier, N. R. F., *Studies of abnormal behavior in the rat.* I. The neurotic pattern and an analysis of the situation which produces it, New York: 1939, Harper and Bros., pp. 81.
13. Patrick, J. R., Studies in rational behavior and emotional excitement. II. The effect of emotional excitement on rational behavior in human subjects, *J. comp. Psychol.*, 1934, **18**, 153–195.
14. Sanders, M. J., An experimental demonstration of regression in the rat, *J. exper. Psychol.*, 1937, **21**, 493–510.
15. Spence, K. W., The nature of discrimination learning, *Psychol. Rev.*, 1936, **43**, 427–449.

VARIABILITY IN LEARNING UNDER MASSED AND SPACED PRACTICE *

BY HENRY E. GARRETT,

Columbia University

I. Problem and Setting

This paper is concerned with the relative economy of massed or concentrated practice as compared with spaced or distributed practice. Under these two conditions of learning, the effect of practice periods of different lengths upon the same task and upon three tasks differing in difficulty was determined. An attempt has been made, also, to determine the relative strengths of short intensive practice and of individual differences in so far as these contribute to total subject variability throughout the learning period.

The problem of massed *versus* spaced learning is not a new one, and much work has been done on this topic. Recent reviews of the literature (**2**) (**18**) reveal that most investigators have found practice distributed over several sittings more economical of learning time than practice concentrated into a single period. This general result holds whether the tasks to be learned involve verbal materials (*e.g.*, nonsense syllables, poetry, prose, etc.) or whether they involve the acquisition of those motor or mechanical skills necessary for success in such activities as typing, archery, or mirror drawing. Similarly, it has been shown that in animal learning distributed practice has the advantage over massed practice. Rats, for example, learn mazes, puzzles, and discrimination habits better when practice is spaced over several intervals.

Despite the agreement in this field, it is nevertheless difficult to draw general conclusions because of the numerous variations in experimental set-up and the large number of variables involved. The latter include—among others—the

* This is one of a series of studies prepared under the auspices of the Columbia University Council for Research in the Social Sciences.

character of the task to be learned and its degree of difficulty; the number and length of the practice periods; the length of the intervals between trials; and the condition of the subjects (*e.g.*, their previous training) when the experiment was begun. Important also are the age and sex of the subjects and their motivation and interest. Such technical and statistical questions as the equality of units employed in measuring improvement, the reliability and validity of measuring instruments, and the methods used in gauging improvement must also be considered (**1**) (**12**) (**13**). The number of permutations and combinations of these experimental variables necessary before any very general result can be reached is well-nigh inexhaustible.

It goes without saying that in the present experiment we have not examined all of the experimental possibilities listed above. We have, however, considered some of the more important factors. Also, through the analysis of variance technique we have been able to attack the question of the relative contributions to group variability of practice and individual differences under various conditions.

II. Procedure

1. *Subjects*

The subjects in this experiment were 96 men students of the College of the City of New York and 78 women students in Barnard College. Most of these subjects were sophomores or juniors and all were students in general psychology. The age range in the group was from 16 to 25 years.

2. *Tests*

Three measures of symbolic or ideational learning were used: a *symbol-digit substitution test*, a *code learning test*, and an *artificial language test*.

(1) *Symbol-Digit Test.*—This is a familiar test of learning ability. At the top of the blank are 9 circles, each containing a symbol and a number. Below this key, arranged in two columns, are 200 symbols in groups of five. Spaces are provided in which each symbol may be represented by its appropriate number.

(2) *Code Learning Test.*—This test was adapted from the code placed at Year XVI, Average Adult, in the 1916 Stanford-Binet. At the top of the test page, the code was printed with directions for its use. Two sample sentences were given below the directions. The test proper consisted of 10 sentences planned to be of equal—or nearly equal—difficulty. These sentences were selected from practice books used by students of typewriting; each sentence contained all 26 letters of the alphabet. Vowels and often used consonants were repeated approximately equally often, and all sentences were of nearly the same length.

(3) *Artificial Language Test.*—This test was taken, with considerable modification, from the American Council on Education Psychological Examination. At the top of the test sheet a vocabulary of 19 words (pronouns, demonstrative adjectives, prepositions, conjunctions, and verbs) was given along with a set of six rules for changing nouns and pronouns from singular to plural, indicating past and future tense, and forming nouns, adjectives, and adverbs from verb stems. Three sample sentences illustrated the use of the vocabulary and the application of the rules. The test proper consisted of twelve sentences of almost exactly the same length and difficulty.

3. *Testing Procedure*

The general procedure followed is shown in Table 1. In this table, the numbers 1, 2, 3, 4, designate the different groups of men and women subjects who worked under

TABLE 1

EXPERIMENTAL PROCEDURE

Test	Massed Practice		Spaced Practice	
	5-trials	10-trials	5-trials	10-trials
Symbol-digit	1M 1F	2M 2F	3M 3M	4M 4F
Code learning	1M 1F	2M 2F	3M 3M	4M 4F
Artificial language	1M 1F	2M 2F	3M 3M	4M 4F

the four conditions. 'M' indicates the male subjects, and 'F' the female. Each group of subjects took all three tests under *one* practice condition; and the initial equivalence of the groups (see Table 6) permits a comparison of the relative efficiency of the various practice conditions. Moreover, since results for each practice condition were obtained from both men and women, conclusions regarding sex differences are possible.

Under massed practice the trials followed each other immediately. In spaced practice, one trial on each test was given at experimental periods separated by intervals of approximately three days. In order to make total time spent equal for all experimental conditions, the duration of a single trial under the 5-trial condition was twice that of a trial under the 10-trial condition. This will be clear from Table 2 which gives

TABLE 2

TIME LIMITS FOR 5-TRIAL AND 10-TRIAL PERIODS

Test	Time per Trial	
	5-trial	10-trial
Symbol-digit	2′	1′
Code learning	6′	3′
Artificial language	4′	2′

the time limits for each test under each condition. These time limits were determined by careful preliminary trial with groups of students.

Care was taken to avoid repetition on successive trials of the same part of the test. In the symbol-digit test the subjects were instructed to begin with the *left* column on odd and the *right* column on even trials. In the code learning and artificial language tests subjects began at different designated points in the test on successive trials.

4. *Scoring*

(1) *Symbol-Digit Test.*—The score on this test was the number of correct substitutions made within the time limit.

(2) *Code Learning Test.*—The score here equalled the number of letters correctly written in code. This seemed to be a more adequate way of measuring performance than number of words written, as the unit of the code is a letter, and the words to be written differ considerably in length.

(3) *Artificial Language Test.*—The score was the number of words correctly translated.

III. Evaluation of Tests as to Difficulty

All comparisons—those between spaced and massed practice, between 5- and 10-trial periods, and between the sexes—are made *within* each test and *between* tests. In the latter case, the fact that the tests, though all measuring symbolic learning, differ in degree of difficulty becomes a potentially important factor. It seemed essential before considering results, therefore, that we establish the relative difficulty of our three tests. None of the criteria of difficulty ordinarily used is strictly applicable here. The frequently employed criterion of time required to perform a task cannot be used because all three tests were given by the time-constant method. Even in the work-constant method time is not an adequate means of determining difficulty when there are qualitative differences between the tests. Thus, it might require as long to work out an example in long division as to do a problem in the calculus, but this does not establish the equal difficulty of the two tasks. A second measure of difficulty often used is a comparison of the percentages of subjects passing a test; but this measure can only be applied to tests which can be classified as 'passed' or 'failed' and is useful mainly in scaling the items within a test. A third criterion which suggests itself is the number of items or units done within a specified time: *i.e.*, if an individual can do more units of the one test than of another in the same length of time, the former test is presumably the easier. This criterion, however, assumes equality of the units within and between the tests to be compared—a condition which is not true of the tests here used.

Since a comparison of our tests in terms of absolute units

was not feasible, the following three criteria based upon relative measures were set up:

(*A*) The ratio of the *final* to the *initial* score (F/I) on each test was computed. This ratio is, of course, always greater than unity if any learning has taken place during the period. It should be smaller for an easy test than for a hard test because subjects score comparatively high on the first trial of a simple task, and hence register less improvement than on a difficult task. Our tests were ranked from easy to hard on the basis of this criterion; the easiest test was that with the smallest F/I ratio, and the most difficult that with the largest F/I ratio.

(*B*) The gain from initial to final score on each test was divided by the S.D. of the scores on the initial trial of the test. Expressing gains in σ-terms eliminates the problem of inequality of units and makes possible a direct comparison of gains from test to test. Following the reasoning in (*A*) above, we should expect the σ-gain for an easy test to be smaller than that for a harder test.

(*C*) A third criterion was based upon a consideration of the rate of progress of learning in easy and hard tasks. It seems reasonable to expect that learning in an easy task will be relatively rapid during the early stages of practice and will later taper off as the task becomes more and more mechanical. Because the hard task offers more complex and more numerous associations, it is assumed that improvement will tend to be slow during the first trials and will increase in the latter part of the work period. The gains in the first and second halves of the work periods were determined, therefore, separately. In the 10-trial periods, the gain from trials 1 to 5 and the gain from trials 6 to 10 were calculated; for the 5-trial condition the gain from trial 1 to trial 2 was balanced against the gain from trial 4 to trial 5. The test for which the ratio

$$\frac{\text{Gain in first half of period}}{\text{Gain in second half of period}}$$

was largest was taken to be the easiest task; that for which the ratio was smallest the hardest task. Results obtained from

the application of these criteria to our three tests are shown in Table 3. Rank 1 indicates the easiest test by a given criterion, rank 2 the next easiest, and rank 3, the hardest.

The results are more clear-cut for massed than for spaced practice. For massed practice, the order of difficulty from

TABLE 3

ORDER OF MERIT (RELATIVE DIFFICULTY) OF THREE TASKS IN TERMS OF THREE CRITERIA, C_A, C_B, AND C_C

Practice Condition	Symbol-digit			Code Learning			Artificial Language		
	C_A	C_B	C_C	C_A	C_B	C_C	C_A	C_B	C_C
Massed 10 M.....	2	3	3	1	1	1	3	2	2
Massed 10F.....	2	3	2	1	1	1	3	2	3
Massed 5M.....	2	2	2	1	1	1	3	3	3
Massed 5F.....	2	2	1	1	1	2	3	3	3
Spaced 10M.....	2	1	2	1	3	1	3	2	3
Spaced 10F.....	3	3	2	1	1	1	2	2	3
Spaced 5M.....	3	2	3	1	1	2	2	3	1
Spaced 5F.....	1	2	2	2	1	1	3	3	3

easy to hard is, according to all three criteria, code learning, symbol-digit, artificial language. The same order holds also —though not so consistently—for spaced practice.

IV. RESULTS

1. *Massed vs. Spaced Learning*

The relative efficiency of massed *vs.* spaced learning was evaluated in terms of the gain in score made under each condition. In the 5-trial work periods this gain equalled the increase in mean score from trial 1 to trial 5; in the 10-trial period the gain was calculated from trial 2 to trial 10. In this way gains were measured for equal periods of time practised under the two conditions. The difference between the gains under massed and spaced practice was also determined for each test and for each condition (*i.e.*, for 5-trials and 10-trials, for males and females). All differences were evaluated in terms of their respective standard errors. Table 4 presents a summary of these results.

It is clear that in the easier tests (code learning and symbol-digit) spaced practice was more effective than massed for both

men and women. There is just one exception to this finding: on the symbol-digit test the women subjects gained more in the 5-trial period under massed than under spaced practice—but the difference is not significant. The advantage of spaced learning is greater for men than for women, as shown by the larger critical ratios. In the artificial language test, the most

TABLE 4

DIFFERENCES BETWEEN GAIN IN MASSED AND GAIN IN SPACED LEARNING FOR DIFFERENT NUMBERS OF TRIALS *

(1) *Code Learning Test*

Condition	Sex	*N*	Diff. in Gain	Favor	σGain	Gain Diff. / σGain	
10 Trials.....	M	40	8.67	Spaced	2.75	3.15	Significant
10 Trials.....	F	36	2.96	Spaced	3.27	.91	Not significant
5 Trials.....	M	50	6.79	Spaced	5.27	1.29	Not significant
5 Trials.....	F	37	4.05	Spaced	5.00	.81	Not significant

(2) *Symbol-Digit Test*

Condition	Sex	*N*	Diff. in Gain	Favor	σGain	Gain Diff. / σGain	
10 Trials.....	M	39	6.78	Spaced	2.66	2.55	Significant
10 Trials.....	F	36	1.20	Spaced	2.33	.52	Not significant
5 Trials.....	M	51	18.70	Spaced	3.71	5.04	Significant
5 Trials.....	F	38	7.25	Massed	4.08	1.78	Not significant

(3) *Artificial Language Test*

Condition	Sex	*N*	Diff. in Gain	Favor	σGain	Gain Diff. / σGain	
10 Trials.....	M	39	2.27	Spaced	1.28	1.77	Not significant
10 Trials.....	F	36	1.75	Massed	1.32	1.33	Not significant
5 Trials.....	M	50	12.41	Massed	1.80	6.89	Significant
5 Trials.....	F	38	5.35	Massed	2.26	2.37	Significant

* A critical ratio of 2.00 or more was considered significant.

difficult of the three, massed practice is superior to spaced in three out of four comparisons (two are significant).

One explanation of these findings which immediately suggests itself is that, in the symbol-digit and code tests, those groups working under massed practice were initially better than those working under spaced practice; and hence the latter made the greater gains. Conversely, in the artificial language

test, if the groups working under spaced practice were initially better than those working under massed practice, the latter groups might be expected to gain more. As shown in Table 5 neither of these hypotheses is true. The two groups compared differed significantly in only one instance (5-trials, men)

TABLE 5

INITIAL SCORES OF MASSED AND SPACED PRACTICE GROUPS IN THREE TESTS

(1) *Code Learning Test*

Condition	Sex	Initial Score		Diff.	$\sigma_{Diff.}$	$\frac{Diff.}{\sigma_{Diff.}}$
		Massed	Spaced			
10 Trials....	M	55.85	58.00	2.15	2.93	.73
10 Trials....	F	58.29	59.21	.92	4.89	.19
5 Trials....	M	110.00	122.42	12.42	5.62	2.21
5 Trials....	F	118.20	106.26	11.94	9.50	1.26

(2) *Symbol-Digit Test*

Condition	Sex	Initial Score		Diff.	$\sigma_{Diff.}$	$\frac{Diff.}{\sigma_{Diff.}}$
		Massed	Spaced			
10 Trials....	M	34.53	36.23	1.70	2.16	.79
10 Trials....	F	36.79	33.79	3.00	2.38	1.26
5 Trials....	M	62.62	61.53	1.09	4.11	.27
5 Trials....	F	66.70	66.90	.20	4.97	.04

(3) *Artificial Language Test*

Condition	Sex	Initial Score		Diff.	$\sigma_{Diff.}$	$\frac{Diff.}{\sigma_{Diff.}}$
		Massed	Spaced			
10 Trials....	M	11.21	12.00	.79	1.29	.61
10 Trials....	F	17.79	16.75	1.04	1.41	.74
5 Trials....	M	20.06	19.74	.32	1.32	.24
5 Trials....	F	29.30	29.15	.15	2.79	.05

and here the greater gain was made by the initially superior (spaced practice) group. There is no *a priori* reason, therefore, for expecting one group to profit more by practice than the other.

The finding that spaced practice favors learning an easy task while massed practice is more effective for difficult tasks is contrary to the result usually reported (**18**). In our results, the reason is probably to be found in the fact that in our easy

tests the associations were simple and discrete: a symbol to be connected with a letter, or a letter with a sign. These connections were readily acquired by our subjects and the chief function of practice was to increase the subjects' speed in the task. For this purpose spaced practice apparently has an advantage over massed (**5**). However, the artificial language test offered all of the problems of learning a new language. In this more complex task each word had to be classified as to part of speech, its meaning learned, and several 'rules' applied. Time was needed for subjects to acquire tools with which to work. Hence it seems reasonable that continued application rather than short, interrupted practice should prove to be superior.

2. *Five-Trial vs. Ten-Trial Periods*

Two points of method are especially relevant to the comparison of 5-trial and 10-trial periods and are, therefore, repeated here. In the first place, each trial in a 5-trial period was twice as long as each trial in a 10-trial period; hence, the total working time in the two periods was equalized. Secondly, learning in the 5-trial periods was measured by the gain in score from trial 1 to trial 5, while learning in the 10-trial periods was calculated from trial 2 to trial 10. Gain was thus determined for equal periods of time and from a constant initial practice level. Hence, any differences in amount learned should reveal whether it is more economical to distribute learning over a series of short, fairly numerous, trials or over a series of longer and fewer trials. Table 6 presents the difference in amount gained by the 5-trial as against the 10-trial groups.

The advantage is overwhelmingly in favor of the 5-trial period. In all cases but one there is a significant difference favoring the longer, less frequent, trial period. This is as true for massed as for spaced learning; for easy as for hard tasks. Several factors may serve to explain this result. One advantage which may be ascribed to the longer trial is that it gives the subject a greater opportunity to become familiar with the task. Secondly, his set is not repeatedly interrupted as in the shorter period where he is put under the necessity

each time of 'getting back to work.' It is probably true, too, that for the performance of any task there exists an 'optimum' work period, the discovery of which requires experimentation with a graduated scale of practice periods of different lengths. This problem of the optimum response time has been investigated in reaction-time studies. It has been experimentally

TABLE 6

DIFFERENCES BETWEEN GAIN IN 5-TRIAL PERIODS AND GAIN IN 10-TRIAL PERIODS, FOR BOTH MASSED AND SPACED LEARNING

(1) *Code Learning Test*

Condition	Sex	*N*	Diff.	Favor	$\sigma_{Diff.}$	$\frac{Diff.}{\sigma_{Diff.}}$
Massed.....	M	51	18.85	5 Trials	3.98	4.74
Massed.....	F	32	30.19	5 Trials	4.10	7.36
Spaced.....	M	39	16.97	5 Trials	4.42	4.41
Spaced.....	F	41	31.28	5 Trials	4.35	7.19

(2) *Symbol-Digit Test*

Condition	Sex	*N*	Diff.	Favor	$\sigma_{Diff.}$	$\frac{Diff.}{\sigma_{Diff.}}$
Massed.....	M	51	11.35	5 Trials	2.39	4.75
Massed.....	F	32	18.02	5 Trials	3.46	5.21
Spaced.....	M	39	23.27	5 Trials	3.89	5.98
Spaced.....	F	42	9.57	5 Trials	3.17	3.02

(3) *Artificial Language Test*

Condition	Sex	*N*	Diff.	Favor	$\sigma_{Diff.}$	$\frac{Diff.}{\sigma_{Diff.}}$
Massed.....	M	50	11.67	5 Trials	1.42	8.22
Massed.....	F	32	10.69	5 Trials	1.78	6.01
Spaced.....	M	39	3.01	10 Trials	1.69	1.78
Spaced.....	F	42	7.09	5 Trials	1.92	3.69

established, for instance, that a foreperiod of about 2″ gives better results (shorter reactions) than longer or shorter periods (**4**) (**15**) (**17**). Similarly, Garrett (**9**) has shown that there exists an 'optimal' response time for judging differences among weights, linear magnitudes, and handwriting specimens. In the present study, apparently, the longer work periods approximate the unknown 'optimum' more closely than do the shorter periods.

3. *Sex Differences*

A comparison of the relative learning efficiency of men and women under each condition of work is shown in Table 7. Within each test the sex differences are fairly consistent. In code learning, three of the four differences favor women; in the symbol-digit test the men have the advantage in three of the

TABLE 7

DIFFERENCES BETWEEN GAINS MADE BY MEN AND WOMEN IN EACH TEST UNDER MASSED AND SPACED PRACTICE

(1) *Code Learning Test*

Condition	*N*	Diff.	Favor	$\sigma_{Diff.}$	$\frac{Diff.}{\sigma_{Diff.}}$
Massed 10 trials . . .	32	2.56	Women	3.07	.83
Spaced 10 trials . . .	44	3.15	Men	2.97	1.06
Massed 5 trials . . .	51	13.90	Women	4.82	2.88
Spaced 5 trials . . .	36	11.16	Women	5.44	2.05

(2) *Symbol-Digit Test*

Condition	*N*	Diff.	Favor	$\sigma_{Diff.}$	$\frac{Diff.}{\sigma_{Diff.}}$
Massed 10 trials . . .	31	.07	Men	2.36	.03
Spaced 10 trials . . .	44	5.65	Men	2.64	2.14
Massed 5 trials . . .	52	6.60	Women	3.49	1.89
Spaced 5 trials . . .	37	19.35	Men	4.27	4.53

(3) *Artificial Language Test*

Condition	*N*	Diff.	Favor	$\sigma_{Diff.}$	$\frac{Diff.}{\sigma_{Diff.}}$
Massed 10 trials . . .	31	1.21	Women	1.32	.92
Spaced 10 trials . . .	44	2.81	Men	1.28	2.20
Massed 5 trials . . .	51	.23	Women	1.86	.12
Spaced 5 trials . . .	37	7.29	Women	2.21	3.30

four comparisons; while in the artificial language test the women surpass in three out of four cases. Although the differences in six cases do not reach statistical significance, their size and consistency of direction are striking.

The relative superiority of men or of women in learning of this sort is independent of the work condition, and depends rather upon the nature of the material to be learned. The women surpass the men on the two tests concerned with

language, while the men are superior in learning the less verbal (numerical-spatial) connections of the symbol-digit test. This finding is in line with the results of other studies in which women have been found to be superior in handling 'verbal' material, and men in dealing with 'mathematical' and 'spatial' material (3) (6) (10) (11).

4. *Practice and Variability*

In the preceding sections the relative efficiency of several practice conditions has been discussed. It remains now to examine more closely the variability of our groups under practice conditions. Two questions arise: (1) Did practice tend to make the groups more homogeneous, more variable, or leave them unchanged? (2) Which of the two factors causing individual differences is the more effective: sample variability or practice variability?

To answer our first question, each group of subjects was divided into two groups: those whose scores on the initial trial fell *above* the mean of the entire group, and those whose scores fell *below* the general mean. For each of these sub-groups the mean gain was determined and the difference between the two means obtained. This difference was divided by its standard error to give the critical ratio or *t*-value; and from the table of *t* (7) p. 166), the corresponding probability (*P*) was read. Table 8 presents these data.

Neither of the sub-groups has a consistent advantage over the other in amount gained. Differences are as often in favor of the lower scoring sub-groups as they are in favor of the higher scoring sub-groups; and none of the differences is significant. The values of *P* all exceed .05, ordinarily taken as the boundary value for significance. This finding is in accord with other studies of this sort (1) (12). Variability under continued practice exhibits neither convergence nor divergence; instead each subject's position in the group is consistently maintained.

In order to evaluate, next, the contributions of practice and of individual differences to variability throughout the experiment, the total variance was analyzed. Fisher's method

for two criteria of classification was utilized. As the method of analysis of variance has not been widely used in psychology,

TABLE 8

DIFFERENCES BETWEEN GAINS MADE BY HIGHER AND LOWER SCORING SUB-GROUPS

(1) *Code Learning Test*

Condition	Sex	Higher Scoring Sub-Group		Lower Scoring Sub-Group		*D*	*t*	Degrees of Freedom	*P* between
		N	Mean Gain	*N*	Mean Gain				
Massed 10....	M	9	26.2	11	31.5	5.3	1.64	18	.10–.20
Massed 10....	F	7	31.6	7	31.9	.3			
Spaced 10....	M	11	40.0	11	35.7	4.3	.73	20	.40–.50
Spaced 10....	F	13	32.6	11	38.0	5.4	1.04	22	.30–.40
Massed 5....	M	17	44.0	16	51.6	7.6	.74	31	.40–.50
Massed 5....	F	8	60.8	12	62.7	1.9	.19	18	.80–.90
Spaced 5....	M	9	54.4	10	55.1	.7			
Spaced 5....	F	9	63.2	10	68.4	5.2	.47	17	.60–.70

(2) *Symbol-Digit Test*

Condition	Sex	Higher Scoring Sub-Group		Lower Scoring Sub-Group		*D*	*t*	Degrees of Freedom	*P* between
		N	Mean Gain	*N*	Mean Gain				
Massed 10....	M	10	22.1	9	28.2	6.1	1.44	17	.10–.20
Massed 10....	F	7	21.3	7	28.6	7.3	1.88	12	.05–.10
Spaced 10....	M	10	33.3	12	30.0	3.3	.61	20	.50–.60
Spaced 10....	F	11	28.2	13	24.4	3.8	.80	22	.40–.50
Massed 5....	M	15	34.3	19	38.0	3.7	.88	32	.30–.40
Massed 5....	F	8	46.0	12	40.9	5.1	.55	18	.50–.60
Spaced 5....	M	12	50.5	7	49.3	1.2	.12	17	.90–.95
Spaced 5....	F	9	34.6	11	36.6	2.0	.28	18	.70–.80

(3) *Artificial Language Test*

Condition	Sex	Higher Scoring Sub-Group		Lower Scoring Sub-Group		*D*	*t*	Degrees of Freedom	*P* between
		N	Mean Gain	*N*	Mean Gain				
Massed 10....	M	10	14.5	9	15.6	1.1	.39	17	.70–.80
Massed 10....	F	6	16.8	8	15.8	1.0	.40	12	.60–.70
Spaced 10....	M	9	17.9	13	16.8	1.1	.64	20	.50–.60
Spaced 10....	F	12	15.5	12	14.4	1.1	.55	22	.50–.60
Massed 5....	M	14	27.9	19	25.2	2.7	.82	31	.40–.50
Massed 5....	F	8	23.0	12	29.5	6.5	1.87	18	.05–.10
Spaced 5....	M	10	16.0	9	12.3	3.7	.94	17	.30–.40
Spaced 5....	F	8	23.1	12	20.5	2.6	.59	18	.50–.60

a simple illustration of the application of the method when there are two criteria of classification is given in Table 9. The illustration is divided into three parts, *A*, *B*, and *C*. The data, which are artificial, deal with the problem of practice and

TABLE 9

EXAMPLE TO SHOW THE APPLICATION OF THE METHOD OF ANALYSIS OF VARIANCE TO A PROBLEM IN WHICH THERE ARE TWO CRITERIA OF CLASSIFICATION (THE DATA ARE ARTIFICIAL)

(A) All Variance Attributable to Subjects; no Interaction

Subjects	First Trial	Second Trial	Sum
A	10	10	20
B	15	15	30
C	20	20	40
D	25	25	50
E	30	30	60
	100	100	200

1. Correction = $(200)^2/10 = 4000$
2. Total Sum of Squares = $4500 - 4000 = 500$
3. Between *M*s of Trials = $1000 - 1000 = 0$
4. Between Subjects' *M*s = $\frac{20^2 + 30^2 + 40^2 + 50^2 + 60^2}{2} = 4500 - 4000 = 0$

	DF	Sum of Squares
Between Trials	1	0
Between Subjects	4	500
Interaction	4	0
Total	9	500

(B) All Variance Attributable to Practice; no Interaction

Subjects	First Trial	Second Trial	Sum
A	10	20	30
B	10	20	30
C	10	20	30
D	10	20	30
E	10	20	30
	50	100	150

1. Correction = $(150)^2/10 = 2250$
2. Total Sum of Squares = $2500 - 2250 = 250$
3. Between *M*s of Trials = $2500 - 2250 = 250$
4. Between Subjects' *M*s = $2250 - 2250 = 0$

	DF	Sum of Squares
Between Trials	1	250
Between Subjects	4	0
Interaction	4	0
	9	250

TABLE 9—*Continued*

(*C*) *All Variance Interaction; no Practice or Subject Variance*

Subjects	First Trial	Second Trial	Sum
A.....	10	30	40
B.....	15	25	40
C.....	20	20	40
D.....	25	15	40
E.....	30	10	40
	100	100	200

1. Correction = $(200)^2/10 = 4000$
2. Total Sum of Squares = $4500 - 4000 = 500$
3. Between *M*s of Trials = $4000 - 4000 = 0$
4. Between Subjects' *M*s = $4000 - 4000 = 0$

	DF	Sum of Squares
Between Trials.....	1	0
Between Subjects.....	4	0
Interaction.....	4	500
	9	500

variability treated in this paper. In example *A*, five subjects have taken a test *twice*—scores on the second test are identical with those on the first. All of the variance, therefore, is attributable to *subjects*: there is no practice variance since there is no difference in mean from trial 1 to trial 2. Also, there is no *interaction* between subjects and trials as each subject scored in trial 2 exactly as he scored in trial 1. In example *B*, each subject improves his initial score by 10 points in trial 2. Here all of the variance must be attributed to *practice*. There is no subject variability since all of the subjects scored exactly alike in trial 1, and exactly alike in trial 2, so that subject variability is zero on both trials. Again there is no interaction of subjects and trials, as all subjects reacted alike to a given trial. Finally, in example *C*, all of the variance is interaction: differential action of practice upon subjects. There is no difference between the means and S.D.'s on the two trials and hence no practice variance. Furthermore, there is no difference between the mean scores of the subjects on the two trials (the sum of the two scores for each is 40) and hence there is no subject variability.

In actual problems in which there are two criteria of classification the total variance will, ordinarily, be made up of all

three factors—practice variance, subjects' variance, and interaction. Interaction is our most valid estimate of experimental error (**8**, pp. 101–104). In the present problem, the practice variance, the sample variance, and the interaction

TABLE 10

F-VALUES FOR PRACTICE AND FOR SUBJECTS IN THREE TESTS

(1) *Code Learning Test*

Condition	Sex	Practice *F*	$F_{(P = .01)}$	Subjects *F*	$F_{(P = .01)}$
Massed 10	M	89.51	2.52	44.54	2.02
Massed 10	F	50.30	2.57	73.89	2.29
Massed 5	M	65.02	3.47	12.74	1.83
Massed 5	F	115.57	3.58	51.91	2.16
Spaced 10	M	98.86	2.50	26.92	1.95
Spaced 10	F	98.69	2.50	33.59	1.89
Spaced 5	M	64.93	3.60	21.96	2.21
Spaced 5	F	95.33	3.60	42.04	2.21

(2) *Symbol-Digit Test*

Condition	Sex	Practice *F*	$F_{(P = .01)}$	Subjects *F*	$F_{(P = .01)}$
Massed 10	M	54.03	2.53	15.17	2.06
Massed 10	F	70.88	2.56	47.25	2.28
Massed 5	M	141.55	3.47	19.78	1.82
Massed 5	F	70.18	3.58	22.94	2.15
Spaced 10	M	66.58	2.50	29.67	1.97
Spaced 10	F	81.85	2.50	38.95	1.88
Spaced 5	M	72.91	3.60	21.12	2.21
Spaced 5	F	55.72	3.58	21.65	2.22

(3) *Artificial Language Test*

Condition	Sex	Practice *F*	$F_{(P = .01)}$	Subjects *F*	$F_{(P = .01)}$
Massed 10	M	82.64	2.53	26.94	2.06
Massed 10	F	70.59	2.56	51.06	2.28
Massed 5	M	200.81	3.47	12.67	1.83
Massed 5	F	115.61	3.58	19.75	2.16
Spaced 10	M	96.17	2.50	24.56	1.97
Spaced 10	F	61.07	2.50	35.07	1.90
Spaced 5	M	40.89	3.60	9.55	2.21
Spaced 5	F	61.80	3.58	17.63	2.22

have been computed separately for each of the eight experimental conditions under which each test was administered.

Practice and subject variance were divided in turn by the interaction to give the *F*-ratios shown in Table 10. From the *F*-table in Snedecor (**14**) the critical value of *F* corresponding

to a P of .01 for the appropriate degrees of freedom was determined. Since the obtained values of F far exceed these .01 points, *both* practice and sample variability contribute significantly to the total variance in all three tests. The F's are larger for practice, however, than for subjects. Hence, it is clear that, in general, practice contributed more than sample variability to the total.

In order to determine more precisely the *relative* strength of our two factors (practice and sample) the ratio of their mean squares was next calculated: the mean square for practice was divided by the mean square for subjects. This new F was evaluated in terms of the F for the .01 point of P corresponding to the appropriate degrees of freedom in the two mean squares. If the calculated F-ratio exceeded the value of $P = .01$, it was concluded that practice contributed significantly more to total variance than did subject-differences. Table 11 gives these data.

In all cases but one, practice contributes more to total variance throughout the experiment than individual differences. In most cases the advantage is not significant, but it is interesting to find it so consistent. Suppose we assume, tentatively to be sure, that, for practice variability constant, differences in score represent native differences in ability. Then it follows that training (intensive practice) contributes more to variability throughout the experiment than native differences (**16**). It could be objected, of course, that subject variability represents previous training as well as native ability; and that practice effect is conditioned by native aptitude; in other words that neither factor represents a clear-cut variable. While this is true to a degree unknown, the novel character of our tests and the common background and training of our subjects make it probable that sample variability reflects native differences to a very large degree. Also, 'training,' *i.e.*, opportunity to practice, was constant for all subjects. With certain qualifications, therefore, we conclude that in our experiments training (defined as intensive practice) was more effective than native differences in making for variability.

It should be made clear that even if the conclusion just stated is more nearly true than it probably is, we still have nothing to say on the general problem of heredity and en-

TABLE 11

THE RELATIVE STRENGTH OF PRACTICE AND INDIVIDUAL DIFFERENCES IN THEIR CONTRIBUTION TO TOTAL VARIANCE

(1) *Code Learning Test*

Condition	Sex	F (Practice/Subjects)	F (P = .01)
Massed 10	M	2.01	3.52
Massed 10	F	1.47	5.11
Massed 5	M	5.10	3.97
Massed 5	F	2.23	4.50
Spaced 10	M	3.67	3.40
Spaced 10	F	2.94	3.30
Spaced 5	M	2.96	4.58
Spaced 5	F	2.27	4.58

(2) *Symbol-Digit Test*

Condition	Sex	F (Practice/Subjects)	F (P = .01)
Massed 10	M	3.60	3.60
Massed 10	F	1.51	4.19
Massed 5	M	7.15	3.95
Massed 5	F	3.07	4.50
Spaced 10	M	2.24	3.40
Spaced 10	F	2.10	3.30
Spaced 5	M	3.45	4.58
Spaced 5	F	2.58	4.50

(3) *Artificial Language Test*

Condition	Sex	F (Practice/Subjects)	F (P = .01)
Massed 10	M	3.07	3.60
Massed 10	F	1.38	4.19
Massed 5	M	15.81	3.97
Massed 5	F	5.87	4.50
Spaced 10	M	3.91	3.40
Spaced 10	F	1.74	3.30
Spaced 5	M	4.28	4.58
Spaced 5	F	3.51	4.50

vironment. Our subjects were highly selected college students; and the training was short, intensive, and upon novel tasks. That such training should influence total variance more than differences in native ability is to be expected in view of the decided homogeneity of our groups. If the groups

had been initially more variable, it is likely that—under the given practice conditions—subject variance would have exceeded practice variance.

Practice is most effective for men and for massed trials. As may be seen in Table 11, in the seven cases (in 24) in which

TABLE 12

SAMPLE AND PRACTICE VARIABILITY OF MALES *vs.* FEMALES

(1) *Code Learning Test*

Condition	Males		Females	
	Practice *F*	Subjects *F*	Practice *F*	Subjects *F*
Massed 10	89.51	44.54	50.30	73.89
Massed 5	65.02	12.74	115.57	51.91
Spaced 10	98.86	26.92	98.69	33.59
Spaced 5	64.93	21.96	95.33	42.04

(2) *Symbol-Digit Test*

Condition	Males		Females	
	Practice *F*	Subjects *F*	Practice *F*	Subjects *F*
Massed 10	54.03	15.17	70.88	47.25
Massed 5	141.55	19.78	70.18	22.94
Spaced 10	66.58	29.67	81.85	38.95
Spaced 5	72.91	21.12	55.72	21.65

(3) *Artificial Language Test*

Condition	Males		Females	
	Practice *F*	Subjects *F*	Practice *F*	Subjects *F*
Massed 10	82.64	26.94	70.59	51.06
Massed 5	200.81	12.67	115.61	19.75
Spaced 10	96.17	24.56	61.07	35.07
Spaced 5	40.89	9.55	61.80	17.63

practice variance is significantly greater than subject variance, five of the seven are for massed practice and six of the seven are for men. This sex difference may be due to one or both of two related causes: (1) the effect of practice may be relatively greater for the men than for the women; (2) subject variability may be greater among the women than among the men. The latter explanation seems to be the more probable. In Table 12 the practice *F*'s for women are as often larger than

those for men as *vice versa*; but the subject *F*'s are, in every case, larger for women than for men. The relative contribution made by individual differences or sample variance to the total variance tends, therefore, to be larger for women than for men. No consistent differences among the tests or among the various conditions appear.

V. Summary

Three tasks differing in difficulty were administered to 174 subjects (college students) under conditions of massed and spaced practice. Five-trial and ten-trial work periods—covering the same total working time—were employed. The following results were obtained:

1. Spaced practice is more efficient than massed practice for the easier tests. Massed practice gave better results for the harder test (see p. 556).

2. For the same working time, 5-trial periods are more effective than 10-trial periods. This result holds for massed and for spaced learning; for hard and for easy tasks.

3. Women are superior to men in tests requiring language associations; men are superior on tests demanding numerical and spatial associations. Material is more important than difficulty or practice conditions in making for sex differences.

4. The variability of our groups remained approximately constant throughout the total working period.

5. Both sample variance and practice variance contribute significantly to total variance over the work period. Of the two factors, practice was more effective than differences among individuals in making for variability over the whole experiment.

(Manuscript received December 16, 1939)

REFERENCES

1. Anastasi, A., Practice and variability. A study in psychological method, *Psychol. Monogr.*, 1934, **45**, pp. 55.
2. Bills, A. G., *General Experimental Psychology*, New York: Longmans, Green and Co., 1934, 216–221.
3. Boynton, P., Sex differences, *Psychol. Bull.*, 1926, **23**, 624–625.
4. Breitweiser, J. V., Attention and movement in reaction time, *Arch. Psychol.*, 1911, **18**, pp. 49.

5. Cook, T. W., Massed and distributed practice in puzzle-solving, *Psychol. Rev.*, 1934, **41**, 330–355.
6. Crawford, C. C., Some comparisons of freshmen boys and girls, *Sch. & Soc.*, 1926, **24**, 494–496.
7. Fisher, R. A., *Statistical Methods for Research Workers*, London: Oliver and Boyd, 1936, 6th edition, xii + 339.
8. Fisher, R. A., *Design of Experiments*, London: Oliver and Boyd, 1935, pp. 252.
9. Garrett, H. E., A study of the relation of accuracy to speed, *Arch. Psychol.*, 1922, **56**, pp. 104.
10. King, I., and M'Crory, J., Freshmen tests at the State University of Iowa, *J. educ. Psychol.*, 1918, **9**, 32–46.
11. Paterson, D. G., and Langlie, T. A., The influence of sex on scholarship ratings, *Educ. Admin. and Superv.*, 1926, 458–468.
12. Perl, R. E., The effect of practice upon individual differences, *Arch. Psychol.*, 1933, **159**, pp. 54.
13. Reed, H. B., The influence of training on changes in variability in achievement, *Psychol. Monogr.*, 1931, **41**, pp. 59.
14. Snedecor, G. W., *Calculation and Interpretation of Analysis of Variance and Covariance*, Ames, Iowa: Collegiate Press, 1934, Table XXXV.
15. Telford, C. W., The refractory phase of voluntary and associative reactions, *J. exper. Psychol.*, 1931, **14**, pp. 36.
16. Thorndike, E. L., Heredity and environment, *J. educ. Psychol.*, 1938, **29**, 161–166.
17. Woodrow, H., Outline as a condition of attention, *J. exper. Psychol.*, 1916, **1**, 23–39.
18. Woodworth, R. S., *Experimental Psychology*, New York: Henry Holt and Co., 1938, 211–216.

EXPERIMENTAL STUDIES IN ROTE-LEARNING THEORY. VI. COMPARISON OF RETENTION FOLLOWING LEARNING TO SAME CRITERION BY MASSED AND DISTRIBUTED PRACTICE

BY CARL IVER HOVLAND

Department of Psychology, Institute of Human Relations, Yale University

Introduction

A number of divergent theories of distribution of practice have as a point in common the explanation that distributed practice is effective because the rest pause between successive trials permits recovery from effects which interfere with learning. Descriptions of the nature of these effects vary greatly from theory to theory; the terms 'interference,' 'decrement,' 'fatigue,' 'conflicting habits' and 'inhibition' have all been applied in this connection. In these formulations performance is assumed to depend upon the strength of the positive learning minus the strength of the interfering process, and the further assumption is made that the interfering tendencies decrease in strength with the passage of time more rapidly than do the positive learning tendencies:

> . . . it is the very *process of forgetting* during the interval which is advantageous, the less fixed and erroneous responsive tendencies being eliminated, while the more fundamental and correct response suffers only a slight loss in retention. At the following period of practice, therefore, the problem is attacked more efficiently because of the loss of these incipient conflicting habits during the preceding interval (**20**, 21–22).

This type of theory might satisfactorily explain more rapid *acquisition* by distributed practice. But following the same logic one should expect better *retention* following *massed* practice, provided the learning by both methods is to the same criterion. This prediction follows because with massed practice considerably more overlearning of the correct responses would be required to overcome the interferences than

with distributed practice where the interpolated rest pauses would prevent the interferences from becoming strongly established. If then during the retention interval 'the less fixed and erroneous responsive tendencies' are eliminated and 'the more fundamental and correct response suffers only a slight loss in retention,' it is apparent that more material should be retained following massed than following distributed practice.

Although this deduction is opposed to educational theory as reflected in the common textbook dictum that 'cramming' is bad for retention, satisfactory evidence concerning this point is not available. That there is some possibility of the predicted experimental outcome is suggested by somewhat analogous results of both Ebbinghaus (**4**) and Woodworth (**21**) which show that long lists which require *more trials* to *learn* are *better retained*. On the other hand are the results of Ebbinghaus indicating that thirty-eight distributed repetitions were better retained on the following day than were sixty-eight continuous readings, and that a list learned for the first time by massed practice was less well retained than following relearning with a day interval between readings. Crucial evidence concerning the relative retention following massed and distributed practice is therefore needed. The present study is devoted to this problem.

The procedure employed in the present study must be distinguished from that used in other distributed practice studies. In the customary distribution of practice experiment (*cf. e.g.* Robinson (**17**), Jost (**14**)) the subjects were given the *same number of trials* by massed and distributed practice and the retention following each compared. In the present experiment retention was compared following learning *to the same criterion* by massed and distributed practice.

Experimental Procedure

The experimental procedure was designed to permit comparison of the retention, at four different intervals, of materials learned to the same criterion by either massed or distributed practice. In order to have each subject act as his own control, eight experimental programs were required. In four, the learning to the criterion of one perfect recitation was by massed practice (six seconds between repetitions); in the other four, by distributed practice (two minutes between repetitions).

Program	Method of Learning	Retention Interval
1	Massed Practice	6 seconds
2	Massed Practice	2 minutes
3	Massed Practice	10 minutes
4	Massed Practice	24 hours
5	Distributed Practice	6 seconds
6	Distributed Practice	2 minutes
7	Distributed Practice	10 minutes
8	Distributed Practice	24 hours

After the appropriate interval, the subjects relearned by massed practice to the original criterion of one perfect recitation. This procedure permitted both recall and relearning scores to be used. Each subject completed two cycles, learning a total of 16 lists in counterbalanced arrangement.

Conditions of the experiment and the general procedure have already been outlined (8). The materials were nonsense syllables learned by the serial anticipation method at a two-second rate of syllable presentation. The instructions forbade rehearsal during the interval between learning and recall. As an added precaution, however, during the two-minute rest pauses between trials during distributed practice, and during the two-minute and the ten-minute intervals between completion of learning and testing of retention, the subjects named automatically presented colors. In the case of the twenty-four hour interval between learning and recall, the subjects named colors for two minutes following learning and for two minutes preceding the tests of retention.

Sixteen bands containing twelve syllables plus a cue symbol were employed for the main experiment. The standard rules for construction of lists were employed. The lists were systematically rotated as previously described (11).

Thirty-two Yale College students were employed as subjects. Subjects learned the lists at the same time every day except Sunday. In the case of the twenty-four hour retention tests, a day elapsed between the recall of the previous list and the learning of the next. Systematic balanced practice orders were prepared in advance and strictly followed. The subjects were unaware of the purpose of the experiment and were not told in advance what procedure would be employed on a given day. Subjects who had been employed in previous experiments had only four practice days. Those serving for the first time were given eight practice series. Practice was given in learning by both massed and distributed trials, and with all four recall intervals.

Results

A comparison of the various experimental programs in terms of the number of trials required to reach the criterion performance is first presented (Table 1). None of the differences between the programs within the massed or within the distributed practice group are statistically significant, but there is a significant difference between the number of trials required to learn by massed practice and the number required to learn by distributed practice. This difference in favor of distributed practice for acquisition is in accord with results reported by earlier investigators.

The principal results of the experiment are those bearing upon the retention, in terms of recall and relearning, following the learning by massed and distributed practice. Table 2 gives the mean number of syllables correctly recalled after varying time intervals following learning to one perfect recitation by massed and distributed practice. It will be observed that at every interval the retention is higher for the materials

TABLE 1

AVERAGE NUMBERS OF TRIALS TO LEARN TO CRITERION OF ONE PERFECT RECITATION WITH EACH EXPERIMENTAL PROGRAM (FOR DESIGNATION OF PROGRAMS *cf.* EXPERIMENTAL PROCEDURE)

Method of Learning	Program	Mean Number of Trials	σ_M
Massed Practice	1	14.13	0.59
	2	14.74	0.61
	3	14.06	0.66
	4	14.93	0.67
Distributed Practice	5	9.80	0.48
	6	9.82	0.54
	7	9.31	0.50
	8	9.22	0.56

learned by distributed practice. The difference is statistically significant at the ten-minute and twenty-four hour intervals (critical ratios are 5.2 and 6.7 respectively).

Another method of comparing retention is by determining the percentage of subjects recalling more following massed and the percentage recalling more following distributed practice at the various time intervals. These results are given in

TABLE 2

AVERAGE NUMBER OF SYLLABLES CORRECTLY RECALLED AT VARIOUS INTERVALS FOLLOWING LEARNING BY MASSED AND DISTRIBUTED PRACTICE (EACH VALUE BASED UPON 64 DETERMINATIONS—2 FOR EACH OF 32 SUBJECTS)

Method of Learning	Interval between Learning and Recall			
	6 Sec.	2 Min.	10 Min.	24 Hrs.
Massed Practice	10.87	11.28	9.33	2.61
σ_M	0.12	0.10	0.20	0.19
Distributed Practice	11.44	11.40	10.48	4.57
σ_M	0.09	0.11	0.18	0.28

Table 3. The same trends indicated in Table 2 are also shown by this method of computation. The present results support those of Patten (**16**) indicating that recall at a ten-minute interval is better following distributed than following massed practice.

It will be observed that in Table 2 recall is lower at every time interval than it was on the criterion trial where a perfect recitation was required. This result is in accord with earlier studies. Ward (**19**) has pointed out that this will occur when learning is carried to a given criterion, since under these cir-

TABLE 3

PERCENTAGE OF SUBJECTS RECALLING MORE SYLLABLES FOLLOWING MASSED OR DISTRIBUTED PRACTICE AT VARIOUS TIME INTERVALS (32 SUBJECTS). STANDARD ERRORS OF PERCENTAGES ARE GIVEN IN PARENTHESES BELOW PERCENTAGES

Interval between Learning and Recall	Percent Recalling More following Distributed Practice	Percent Recalling More following Massed Practice	Percent Recalling Equal Number following Massed and Distributed Practice
6 secs.	43.8 (.087)	21.8 (.073)	34.4 (.084)
2 mins.	37.5 (.086)	25.0 (.077)	37.5 (.086)
10 mins.	46.9 (.088)	15.6 (.065)	37.5 (.086)
24 hrs.	68.8 (.082)	9.4 (.052)	21.8 (.073)

cumstances the peak of a constantly fluctuating performance is systematically used as the criterion and the average performance on the next trial, not being so selected, will usually show a decline. It will also be seen that an initial increase in recall ('reminiscence') is obtained following learning by massed practice, but not following distributed. This confirms a previous study of the writer (**8**). This result, taken in conjunction with the higher level of recall following distributed practice, suggests that reminiscence is the result of a depression of performance at the end of practice, not a true increase with the passage of time.

In Table 4, the retention at various intervals following

TABLE 4

COMPARISON OF PERCENTAGES OF MATERIAL CORRECTLY RECALLED AT VARIOUS INTERVALS FOLLOWING MASSED PRACTICE IN PRESENT EXPERIMENT AND IN STUDIES OF LUH (13) AND WARD (19)

(*Recall at 6 sec. interval* = 100 *percent.*)

	6 Sec.	2 Min.	10 Min.	24 Hrs.
Present study	100	104	86	24
Ward	100	107	90	—
Luh	100	—	—	22

massed practice in the present experiment is compared with the results of other investigators. From the data of Table 2, the percentages of retention in the present experiment have been derived. These results are compared with those of Ward (19) upon immediate, two-minute and ten-minute retention, and those of Luh (13) (recalculated by Ward) upon immediate and twenty-four hour retention. For all comparisons, the base (100 percent) is the immediate recall (six-second retention). It will be observed that the present massed practice results are closely similar to those of Ward and Luh, despite the fact that in their studies the individual letters were recited, while in the present investigation the entire syllable was the unit. No closely comparable data were available for comparison with the present results on retention following distributed practice with a rest interval between each trial.

Further results on retention are given in Table 5. Here are presented the number of trials required to relearn to the

TABLE 5

AVERAGE NUMBER OF TRIALS REQUIRED TO RELEARN TO CRITERION OF ONE PERFECT RECITATION AT VARIOUS INTERVALS FOLLOWING LEARNING BY MASSED AND DISTRIBUTED PRACTICE (2 DETERMINATIONS UPON EACH OF 32 SUBJECTS)

Method of Learning	Interval between Learning and Relearning			
	6 Sec.	2 Min.	10 Min.	24 Hrs.
Massed Practice	0.92	0.48	2.12	6.73
σ_M	0.15	0.12	0.24	0.29
Distributed Practice	0.34	0.36	1.27	4.69
σ_M	0.11	0.14	0.19	0.23
Critical ratio σ	4.4	0.8	3.4	6.3

criterion of one perfect recitation after various time intervals following learning by massed and distributed practice.[1] These results show exactly the same differences between massed and distributed practice as did those on recall scores. At every time interval, the material learned by distributed practice is relearned with fewer trials than that learned by massed, all of the differences except at the two-minute interval being statistically significant. Table 6 confirms these results. This

TABLE 6

PERCENTAGE OF SUBJECTS RELEARNING MORE RAPIDLY WHEN ORIGINAL LEARNING WAS BY MASSED OR DISTRIBUTED PRACTICE (32 SUBJECTS)

Standard errors of percentages are given in parentheses below percentages

Interval between Learning and Relearning	Percent Relearning more Rapidly following Learning by Distributed Practice	Percent Relearning more Rapidly following Learning by Massed Practice	Percent Relearning Equally Rapidly following Learning by Massed and Distributed Practice
6 secs.	43.8 (.087)	31.2 (.082)	25.0 (.077)
2 min.	40.7 (.087)	28.1 (.079)	31.2 (.082)
10 min.	53.2 (.088)	15.6 (.064)	31.2 (.082)
24 hrs.	71.9 (.079)	12.5 (.059)	15.6 (.064)

table gives the percentages of the subjects who relearned with fewer trials by distributed than by massed practice. The advantage of the distributed practice learning becomes especially evident at the longer time intervals.[2] The present results on relearning do not agree with those of Patten (**16**). As an incidental observation he found that relearning after ten minutes was more rapid following massed practice than following distributed. He is unable to account for the divergence of these results from those on recall.

[1] For the present comparisons the absolute numbers of relearning trials are required rather than the more common 'savings scores' since otherwise a different base would be involved for massed and distributed practice.

[2] A recent experiment by Cain and Willey suggests that the divergence between recall following massed and distributed practice becomes even more marked at longer intervals (three and seven days). (Cain, L. F., and Willey, R. de V. The effect of spaced learning on the curve of retention, *J. exper. Psychol.*, 1939, **25**, 209–214.)

Retention curves are plotted in Fig. 1 for recall and Fig. 2 for relearning. The time units are spaced logarithmically. It will be seen that following the first two minutes, retention (both in terms of recall and relearning) decreases almost linearly with the logarithm of time. This is in accord with Strong's results using the recognition method of testing re-

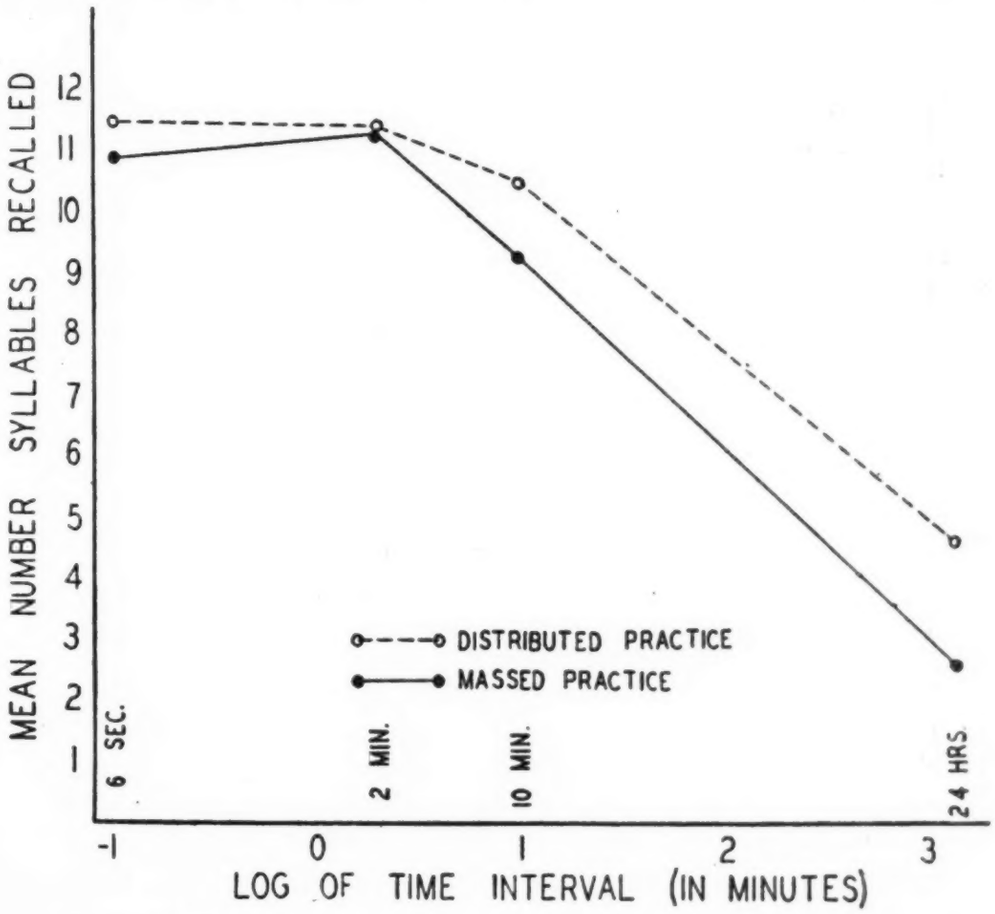

FIG. 1. Mean number of syllables recalled at various time intervals following learning to one perfect recitation by massed and distributed practice. Time intervals are spaced logarithmically. Each point is mean of 64 scores (2 upon each of 32 subjects).

tention (*cf.* Woodworth (**22**)). The initially slow decline in recall might be attributed to the presence of some interfering process which masks the true extent of learning. It is apparent, however, that even with Strong's data, extrapolation on the basis of the logarithmic relationship to a near zero retention interval (*e.g.* 1/1000 sec.) would give a ridiculously high value.

Incidental data are available in the present study with

respect to an apparent divergence between the results of Hull (**11**) and Ward (**19**) as to the relative recall in the various serial positions. Hull showed that after a time interval of 20.5 hours, recall is no better at the ends than at the center of the list. He contrasts these results with the curve of recall at the various serial positions during the comparable stage of original learning. Hull concludes that "whatever is responsi-

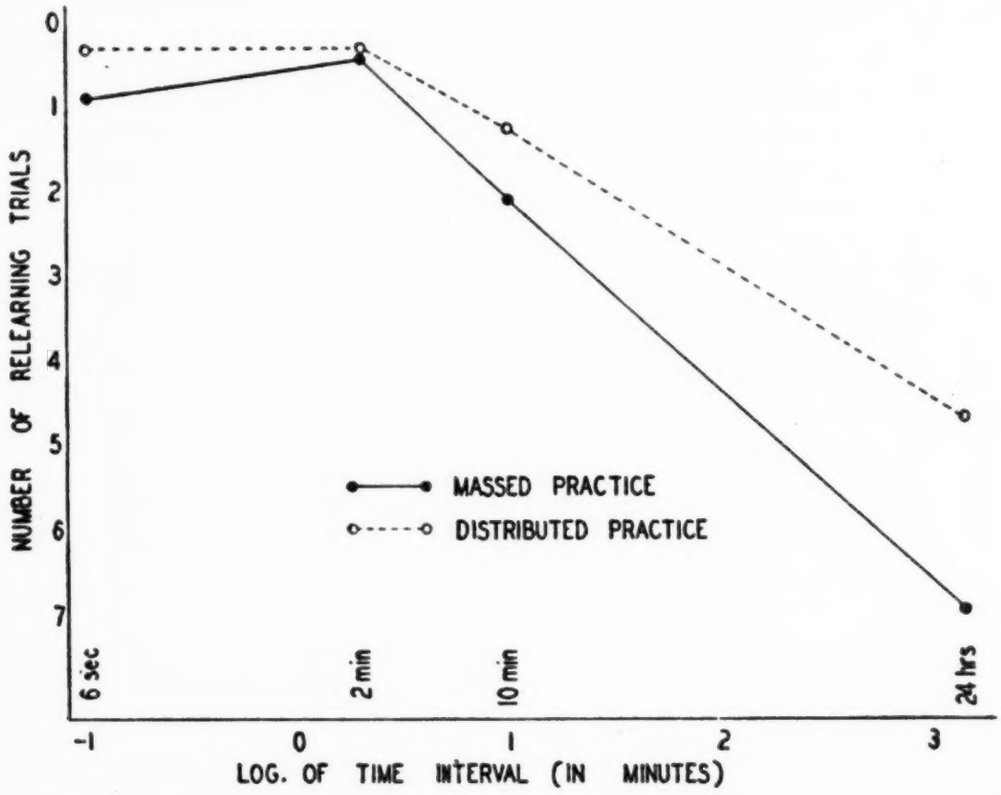

FIG. 2. Mean number of trials required to relearn to criterion of one perfect recitation after various time intervals following learning by massed and distributed practice. Time intervals are spaced logarithmically. Each point is mean of 64 scores (2 upon each of 32 subjects).

ble for the special difficulty in the central portion of the series (or for the special ease at the two ends) has ceased to operate after a period of 20.5 hours" (p. 262). Ward, on the other hand, found pronounced serial position effects with time intervals up to twenty minutes between learning and recall. He states that the better retention of the end syllables is "in rather marked contrast to the findings of Hull." He suggests that the discrepancy is "probably due only to differences in time of testing retention" (p. 37). In Fig. 3, *A*, are given the

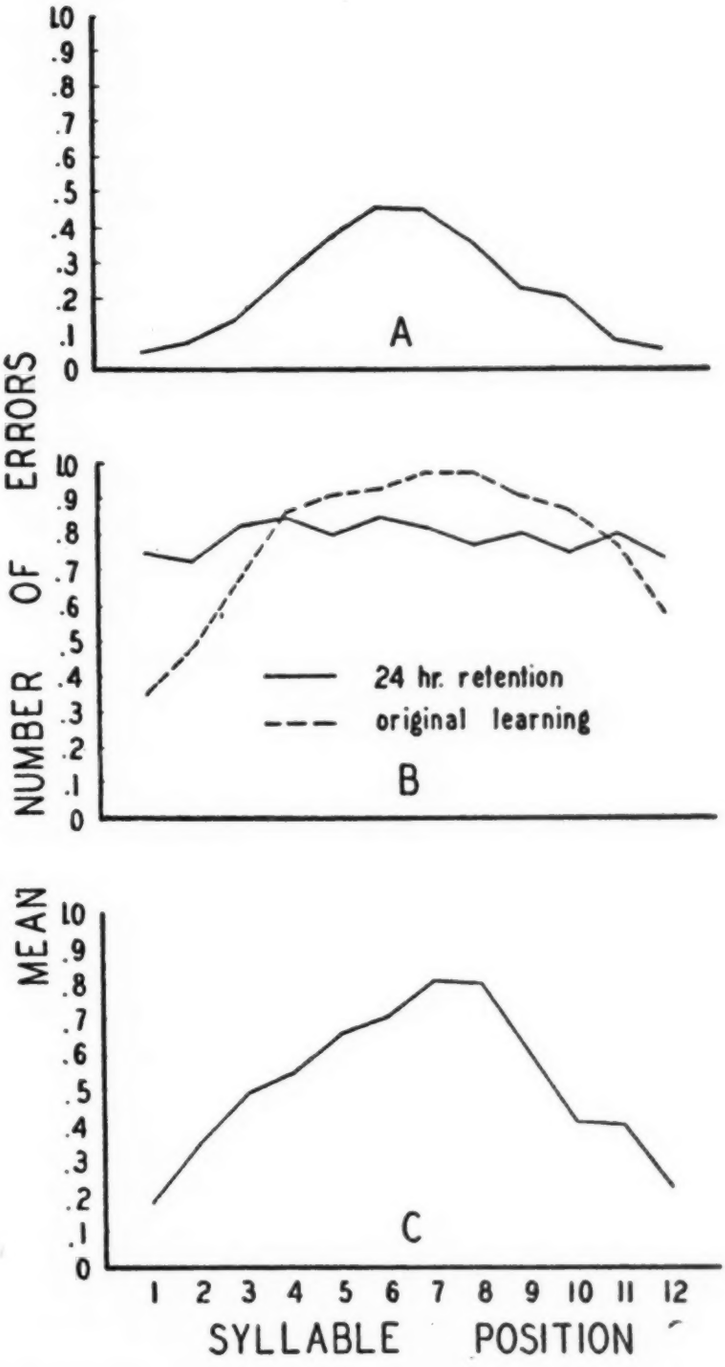

FIG. 3. *A.* Mean number of errors in successive serial positions during recall 10 minutes after learning by massed practice. *B.* Mean number of errors in successive serial positions (1) during first recall trial 24 hours after massed practice learning and (2) during original learning at comparable stage of performance. *C.* Mean number of errors in successive serial positions during second trial of relearning (first trial following recall) 24 hours after massed practice learning. All values based upon mean of 64 cases (2 upon each of 32 subjects).

number of errors in each serial position in the present experiment when recall is ten minutes after massed practice learning. At this interval, the usual serial position effects are clearly evident. The results are closely similar to those of Ward. In the present experiment, the average number of syllables correctly recalled in the end syllable positions (3–4–11–12) at the ten-minute interval was 3.48. At the same interval Ward's subjects recalled 3.29 in these positions. For the central syllables (6–7–8–9) the subjects in the present study recalled 2.49; Ward's, 2.29. The ratio of the difficulty of the central to the end syllables is thus 1.40 in the present experiment, and 1.43 in Ward's, a remarkable agreement in view of differences in experimental procedure. After a twenty-four hour interval, however, no consistent serial differences appear (Fig. 3, *B*). The dotted line in Fig. 3, *B*, shows the errors in the various serial positions at the comparable stage of original learning. The contrast is apparent. These results then indicate that there is no inconsistency between the results of Hull and Ward. At short intervals pronounced differences in the ratio of difficulty of the central to the end syllables are apparent, but long intervals appear to obliterate these differences.

While the curve of errors in the successive syllable positions is relatively flat on the first recall trial after a twenty-four hour interval, the customary greater difficulty of the central syllables is clearly seen on the very next trial. Figure 3, *C*, gives the number of errors in the various serial positions on the relearning trial immediately following the first recall trial at the twenty-four interval. The difference between this curve and that for the first recall trial (Fig. 3, *B*) is evident.

Discussion

The present results showing the superiority of distributed practice over massed for *retention* are in flat contradiction to the deduction from the identical postulates most frequently employed in explaining rapid *acquisition* by distributed practice. It thus appears that either the postulates customarily used are erroneous or additional factors are operative in this situation which require the introduction of new postulates.

The validity of one of the postulates mentioned in the Introduction, namely, that the interferences ('inhibitory effects') decline in strength more rapidly than the positive learning process, has already been called into question, at least in the case of conditioned discrimination, by Hilgard and Humphreys (**7**). These investigators found no significant differences in the retention of responses to positive and negative stimuli after intervals up to 19 months. The specific difficulty raised by their experiment, if found also to obtain in nonsense syllable learning, could, of course, be circumvented by assuming that the positive and interfering processes differ in strength and that this difference, rather than dissimilarity in the nature of the processes, is the significant factor in determining the relative rate of decline. The difference in strength, they point out, could be due either to prior establishment of the positive tendencies, or to greater overlearning. Despite any such modifications, however, the basic disagreement between the predicted and obtained outcome of the present experiment would remain.

Two modifications of the basic postulates suggest themselves as means of bringing the original theory into line with the experimental facts. The first alternative would involve a reformulation of the rate of decline of the interfering processes. In the present systems, such as the one put forward by Hull, Hovland, Ross, Hall, Perkins and Fitch (**12**), it is assumed that the interferences decrease according to the same general function as the positive tendencies, but at a more rapid rate, and that both functions become asymptotic to zero. As an alternative it might be assumed that the interfering process does decrease rapidly at first, but that it then becomes asymptotic to a positive value considerably above zero. In this case inferior retention would be expected following massed practice because the interferences would still be present even after a long period of time. The difference in the initial rate of decline of the two processes could still account for the reminiscence phenomenon. Unfortunately, we have at present so few data by which this possibility could be tested that no decision as to its adequacy can be made, and the formulation is therefore *ad hoc*.

The other possible revision would be the addition of one or more postulates expressing the substance of the Jost laws, *i.e.*, that a new repetition strengthens an older more than a younger association of equal strength, and that older associations decline less rapidly than younger ones of equal strength. This would resolve the present difficulty, since the associations learned by distributed practice have an average age greater than those learned by massed because of the interpolation of rest pauses. Hence they should be better retained, and strengthened to a greater extent by repetition. This alteration would, however, frankly be another attempt to cover the specific difficulty raised by the present study. The writer is of the opinion, moreover, that Jost's laws themselves require explanation, since they do not *explain* the difference in retention as a function of age of association, but merely *describe* it. In other words, it would seem that in a well-integrated theory of learning Jost's laws should be deductions from general principles rather than basic postulates. An explanation of the second law on the basis of the negative acceleration of the forgetting curve would be an attempt in this direction. It may very well turn out, as has frequently happened before, that age is not the fundamental variable, but is only incidental to more basic factors.[3]

An alternative to revising the present postulates would be the adoption of some other theoretical system, in the hope of covering all of the phenomena more satisfactorily. Unfortunately, however, one finds that just as the systems which best explain rapid *acquisition* by distributed practice seem to have no systematic account of better retention following distributed practice, so most of those which suggest satisfactory explanations for better *retention* have difficulty in explaining more rapid acquisition. Gates' theory (**5**), for example, explains superior retention following distributed practice on the basis of overlearning which comes with successive relearning. His illustrative diagram is well known and widely quoted. This type of theory, however, would demand that

[3] *Cf.* McGeoch: "In scientific descriptions of nature time itself is not employed as a causative factor . . . " (**15**, p. 359).

distributed practice require more trials to reach a given criterion of performance, because of the greater overlearning involved, a deduction not in accord with experimental evidence.

Similarly, Guthrie's analysis might successfully predict better retention following distributed practice on the basis of a differential alienation of cues. He might state that during massed practice postures, movements, tensions and the like are more or less continuously present, and hence by conditioning become closely attached to the verbal responses. When practice ceases these cues are rapidly eliminated, since they become incidental accompaniments of other situations, and thus lose their associational connections with the first responses. Rapid forgetting therefore occurs. He might contend, however, that with distributed practice the learning is protracted over a longer period of time so that the postures, movements, tensions and other incidental cues are successively conditioned and alienated. Only the relevant associational cues furnished by the syllables in the list persist from trial to trial. With the passage of time responses learned by distributed practice should, therefore, remain relatively stable, since the incidental cues have been already differentiated and eliminated, and only the direct associational cues remain operative. This theory, however, would likewise demand a longer period for learning by distributed practice to permit the constant substitution of stable associations to take the place of the incidental ones which are being lost, unless additional assumptions are introduced.

Snoddy's (**18**) theory of 'primary and secondary growth' could successfully predict the present results without too great difficulty. It might be assumed that distribution of practice promotes primary growth which is said to be enhanced by interpolated time. Rapid acquisition could be thus explained. Furthermore, primary growth is supposed to be quite stable and this might account for the good retention of material learned with interpolated rest pauses. Massed practice might be assumed to produce both primary and secondary growth. The acquisitions of the latter are assumed to be transient and

unstable, and this might explain their poor retention. It is, of course, the virtue of such a two-process theory that it can 'predict' any phenomenon if it is given freedom to specify the extent of operation of the two opposed factors. Until more information is available, therefore, concerning the situation in which primary and secondary 'growth' are alleged to operate, judgment must be suspended as to the adequacy of this system. A number of vulnerable points in the theory have been pointed out by Doré and Hilgard (**3**).

Another 'dualistic' theory has been advanced by Cook (**1**). He insists that a clear distinction must be made between *intelligent* learning, which consists of the appearance of insight and the discovery of relationships, and *mechanical* learning, which is the fixing and setting of material already understood. On the basis of this dichotomy he makes a number of deductions, including one stating that "massing should be more economical for immediate recall and distribution for delayed recall." While his own data on puzzles do show an advantage of massed practice during the first few trials, experimental results in general have refuted this deduction by showing that distribution is not only economical for delayed recall but also for immediate, at least where the trial is the unit employed in computing economy.

It would seem that a more valuable distinction for our purposes could be made between factors favoring rapid acquisition and those promoting better retention. It is surprising to find how frequently it is assumed that the same factors which might explain rapid acquisition would also explain better recall. One excellent text book (**2**), for example, says under the topic of 'Acquiring' that "in the *learning* of practically any kind of habit, economy is found by spacing the practice with time intervals instead of attempting to get it completely formed all at one sitting. *This is* one of the reasons *why* 'cramming' has always been looked upon as a good enough method to prepare for certain mechanical types of quizzes, perhaps, but surely a bad one for anything like permanent retention." [4]

[4] P. 410. Italics the present writer's.

Factors of the first variety, promoting more rapid *acquisition* by distributed practice, have already been discussed by the writer (**9**, **10**). These are of two general types, similar to those suggested by Doré and Hilgard (**3**). The first group may be labelled 'performance factors.' Here one refers to influences such as fatigue, refractory phase and the like which mask the true extent of learning. Rest pauses may permit recovery from these effects and hence bring about more rapid learning. Opposed to these influences are 'learning factors.' During acquisition both incorrect (and therefore interfering) and correct associations are constantly being established. Distributed practice may be effective because it permits a differential weakening of these two types of associations. If, because of differential strength, the interfering associations weaken more rapidly than the correct associations, a net gain in performance would result from interpolation of rest pauses.

The other phase of the analysis would involve determining the reasons for better *retention* following distributed practice. Why should two lists which can be recited equally well at the present time differ at a later time in their recallability as a result of differences in the method involved in the original acquisition? Again, both performance and learning factors may be involved. The work or performance factors, however, can only result in superior retention of materials learned by massed practice. The passage of time would permit recovery from decremental effects and, since these are always assumed to be greater with massed practice, the materials learned by this method should recover most. Learning factors must therefore be paramount in all situations where superior retention is found following distributed practice. Some of the learning factors we have already discussed. Differences in the average age of associations learned by massed and distributed practice, for example, may be involved. Another difference may exist in the nature of the practice involved. Distributed practice involves periodic recall after a no-practice interval, and each practice trial might therefore be said to be similar to the final recall. If the more nearly the practice approaches the situation in which the material is to be recalled,

the more effective it is, the deduction follows. Lastly, the rate of decline of correct and incorrect associations with the passage of time may be a factor influencing *retention* as well as acquisition. If interferences decline rapidly and completely we should expect superior retention following massed practice, as outlined in the Introduction. If, on the other hand, they decline rapidly at first but then level off at a value above zero, inferior retention following massed practice would be predicted. To cover both reminiscence and distribution of practice phenomena, it might parsimoniously be assumed that the rate of decline of a response is inversely proportional to its strength. Then the recently acquired and weak interferences might decline rapidly while the older and more firmly fixed interferences would be present for a considerable time following cessation of practice. The rapid decline of the weak interferences could account for reminiscence, while the slow decline of the stronger interferences would be able to account for the poor retention following massed practice.

The implication of the above discussion is clear: we cannot make precise predictions about learning phenomena until we know much more than we now do concerning the factors involved. At present, for example, we do not know the extent of operation of 'performance' and 'learning' factors, nor the relative rate of decline of correct and incorrect associations. Extensive research is therefore demanded. One profitable method of approach would be the comparison of reminiscence and distribution of practice phenomena in a wide variety of learning situations, differing in the relative importance of performance and learning factors. Pursuit-rotor learning, nonsense syllable learning and puzzle solving are suggested as three basic types differing in this respect. By systematic comparison of acquisition and retention of such materials valuable information as to the extent of operation of the two types of factors could be obtained. Additional work must also be directed to the problem of retention of correct and incorrect associations at all levels of learning. Only after such data are obtained can accurate predictions be made.

Summary and Conclusions

A comparison was made of the retention of nonsense syllables learned to the same criterion by massed and distributed practice. Thirty-two college students served as subjects, learning twelve syllable lists to a criterion of one perfect recitation by the anticipation method. Six seconds elapsed between trials during massed practice and two minutes between trials during distributed practice. Retention was tested at intervals of six seconds, two minutes, ten minutes and twenty-four hours, but only one interval was used following the learning of a given list. Between distributed practice trials and between learning and recall (with the exception of the twenty-four hour interval) the subjects named colors automatically presented; no rehearsal was thus possible.

The principal results were the following:

1. Recall was better following learning by distributed practice than following learning by massed practice at every time interval despite the fact that fewer trials had been spent in learning by distributed practice. This superiority was reflected in the greater percentage of subjects recalling more after distributed practice learning than after massed practice learning.

2. Similar results were obtained when the numbers of trials required to relearn to the original criterion after each interval were computed: fewer trials were required to relearn at every interval when the original learning had been by distributed practice. This was found to hold not only for averages but for a majority of the individual subjects as well.

3. Forgetting curves were plotted with the time intervals spaced logarithmically. Following the first two minutes, both recall and relearning showed a practically linear decrease as a function of the logarithm of the time. During the first two minutes, however, both curves show a slow rate of forgetting, that following massed practice actually showing an initial increase in retention ('reminiscence'). The apparent divergence between the curve for distributed practice and massed practice learning becomes greater with the passage of time.

4. The discrepancy between the results of Hull and of Ward with respect to serial position effects is shown to be simply a function of the time interval between learning and recall. Pronounced serial position effects are found at the ten minute interval, but after twenty-four hours serial position effects are practically absent. On the very next trial (second relearning trial), however, marked differences in the various serial positions are evident at the twenty-four hour retention interval.

It appears that the major results of the present experiment cannot be explained solely on the basis of a differential rate of decline of the positive learning and of an interfering process—the most commonly accepted theory for distributed practice. It seems necessary to distinguish clearly between factors contributing to more rapid *acquisition* by distributed practice and those contributing to better *retention* following learning by distributed practice. Possible differential factors are suggested.

By assuming that interferences decline at a rate inversely proportional to their strength, the present results can be reconciled with reminiscence data. It would be assumed that both more weak and more strong interferences are built up with massed practice than with distributed. The rapid elimination of the weak interferences would be used to explain reminiscence. The strongly established interferences would be then expected to impair recall for a considerable period following massed practice, so that retention would be superior following distributed practice, where fewer strong interferences are assumed to be established.

(Manuscript received December 16, 1939)

References

1. Cook, T. W., Massed and distributed practice in puzzle solving, *Psychol. Rev.*, 1934, **41**, 330–355.
2. Dashiell, J. F., *Fundamentals of general psychology*, New York: Houghton, Mifflin, 1937.
3. Doré, L. R., and Hilgard, E. R., Spaced practice as a test of Snoddy's two processes in mental growth, *J. exper. Psychol.*, 1938, **23**, 359–374.
4. Ebbinghaus, H., *Über das Gedächtnis*, Leipzig: Duncker and Humblot, 1885.

5. GATES, A. I., *Psychology for students of education* (Rev. ed.), New York: Macmillan, 1930.
6. GUTHRIE, E. R., *The psychology of learning*, New York: Harpers, 1935.
7. HILGARD, E. R., AND HUMPHREYS, L. G., The retention of conditioned discrimination in man, *J. gen. Psychol.*, 1938, **19**, 111–125.
8. HOVLAND, C. I., Experimental studies in rote-learning theory. I. Reminiscence following learning by massed and by distributed practice, *J. exper. Psychol.*, 1938, **22**, 201–224.
9. HOVLAND, C. I., Experimental studies in rote-learning theory. III. Distribution of practice with varying speeds of syllable presentation, *J. exper. Psychol.*, 1938, **23**, 172–190.
10. HOVLAND, C. I., Experimental studies in rote-learning theory. V. Comparison of distribution of practice in serial and paired-associate learning, *J. exper. Psychol.*, 1939, **25**, 622–633.
11. HULL, C. L., The influence of caffeine and other factors on certain phenomena of rote learning, *J. gen. Psychol.*, 1935, **13**, 249–274.
12. HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T., AND FITCH, F., *Mathematico-deductive theory of rote learning—a study in scientific methodology*, New Haven: Yale Univ. Press, 1940.
13. LUH, C. W., The conditions of retention, *Psychol. Monogr.*, 1922, **31**, No. 142.
14. JOST, A., Die Assoziationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen, *Zsch. f. Psychol.*, 1897, **14**, 436–472.
15. MCGEOCH, J. A., Forgetting and the law of disuse, *Psychol. Rev.*, 1932, **39**, 352–370.
16. PATTEN, E. F., The influence of distribution of repetitions on certain rote learning phenomena, *J. Psychol.*, 1938, **5**, 359–374.
17. ROBINSON, E. S., The relative efficiencies of distributed and concentrated study in memorizing, *J. exper. Psychol.*, 1921, **4**, 327–343.
18. SNODDY, G. S., *Evidence for two opposed processes in mental growth*, Lancaster (Pa.): Science Press, 1935.
19. WARD, L. B., Reminiscence and rote learning, *Psychol. Monogr.*, 1937, **49**, No. 220.
20. WARDEN, C. J., The distribution of practice in animal learning, *Comp. Psychol. Monogr.*, 1923, **1**, No. 3.
21. WOODWORTH, R. S., The influence on retention of conditions favoring quickness of learning, *J. Phil.*, 1915, **12**, 246.
22. WOODWORTH, R. S., *Experimental psychology*, New York: Henry Holt, 1938.

'HYPOTHESES' AND 'RANDOM ACTIVITY' DURING THE CONDITIONING OF DOGS[1]

BY W. N. KELLOGG AND IRVIN S. WOLF

Indiana University

I. Introduction

It has been suggested by numerous authors that the concept of trial-and-error learning is inaccurate, if it is meant to imply that the early approach to the learning problem is made up wholly of 'random exploratory movements' or 'chance responses.' The concept of Gestalt insight, the 'means-end-readiness' of Tolman (**8**, **9**) and the 'hypotheses' of Krechevsky (**3**, **4**, **5**) all stand out as exceptions or alternate proposals of what takes place during the learning process. Krechevsky in particular has emphasized the point that the learner's activity breaks itself up into systems of action ('hypotheses') which give a certain coherence to the behavior throughout a number of trials. Although there have been critics of this view (**7**, **10**), the fact has become well established that there are sequences of behavior in the process of learning which are knit together with a consistency beyond the influence of pure chance.

Experimental data related to the problem have been confined, for the most part, to the behavior of rats in the solution of problems which are considered to be at a fairly difficult level for these animals. It is our purpose here to present evidence to show that consistent behavior with sudden or abrupt modifications (which may themselves persist) are to be found even in the flexion conditioning of dogs. And that, if the proper recording techniques are used, it is not necessary to suppose that insight or 'hypotheses behavior' is something

[1] The research work, upon which this publication from the Indiana Conditioning Laboratory is based, was made possible by a series of grants-in-aid of research from the University, and by continual assistance from the National Youth Administration of the United States Government.

which is entirely incompatible with the traditional 'conditioned reflex.'

II. Conditioning Procedure

The animals which were conditioned were placed individually in a small sound-proofed room, in which they could be observed through one-way-vision windows. They stood within a heavy wooden framework or stock which was bolted on top of a table. Each foot was strapped to light balanced levers which permitted a vertical movement of 4 inches. The animal's head was held by means of an adjustable wooden collar, and a loose strap was passed under the abdomen.

The conditioned stimulus was either the sound of a doorbell buzzer or a pure tone of 1000 ~ of 50–55 db. intensity. In either case its duration was 2 sec. The buzz was sounded alone for 1.8 sec. before the presentation of the unconditioned stimulus, and it coincided with the latter for the final .2 sec. The unconditioned stimulus was either an a.c. or a d.c. shock delivered through stainless steel electrodes securely taped to the anterior and posterior surfaces of the foot just above the toes. The intensity of the shock stimulus was in all cases adjusted so as to be just strong enough to elicit an unconditioned flexion reflex of 4 inches. Resistance changes or other conditions within the organism which altered the effectiveness of any particular setting of the stimulus, were immediately counterbalanced by a new adjustment.

The apparatus used for most of the subjects permitted the obtaining of simultaneous graphic records of the movements of all 4 feet, the respiration, and buzz, shock, and time for every trial. It therefore gave a fairly good picture of the behavior of the whole organism. It is chiefly through the technique of recording more than the responses of the conditioned member that the changes in general activity reported in this paper have been brought to light.[2]

III. Observations on 'Random' Behavior

Trial-and-error in Conditioning.—The flexion CR in dogs has usually been described as developing from the general activity of the early trials. Since the dog is necessarily confined in the experimental situation, the most 'random' activity he can possibly display is to lift all 4 feet in rapid succession, throw his body from side to side, bark or howl, and attempt to jerk his head from the restraining collar—in other words, to struggle. At the start of the training, it is said, the animal struggles to the unconditioned shock stimulus. Later he transfers the struggle or random activity to the buzz, and finally he gradually reduces the unnecessary responses of the non-shocked feet, leaving only the localized flexion CR.[3] If

[2] For a more detailed description of the procedure and mechanism of recording, see Kellogg, Davis, and Scott (2).

[3] A review of the literature on this subject has been given elsewhere by the present writers (11).

the responses of the non-shocked feet are now regarded as 'errors,' then their gradual elimination is not unlike the elimination of errors in the typical trial-and-error experiment.

Our own observations have indicated that a trial-and-error description of this nature does not fit a good proportion of the cases, although it holds for many. These observations are supported by the conditioning of nearly 60 animals in a number of different experiments, and include the conditioning of all 4 limbs separately as well as the 'bilateral' or simultaneous conditioning of both front and rear members. The differences which have been found are related in some instances to the lack of randomness of the early behavior; in others to the suddenness and persistence of the modifications which take place.

The 'Random' Struggle Response.—One of the principal objections to the view that the conditioned response is a refinement or selection that works its way out of 'random' struggling, is that some animals simply do not struggle. Subjects which give the neatest and easiest-to-plot conditioning curves are likely to be those which begin by making full unconditioned flexions to the shock stimulus, and which fail, at any time, to add to this the generalized behavior of struggling. There are, on the other hand, hypersensitive individuals that struggle almost continuously at the confinement of the stock or conditioning situation. When in the middle of a struggle spasm an animal of this sort often seems unable to make any reaction at all to the conditioned stimulus. He may support himself by means of the neckpiece and the strap beneath his belly, while all 4 feet are rapidly 'pumping.'

It is not uncommon to find subjects that fail to eliminate struggle even after hundreds of trials. They may become conditioned very well—that is another thing—but they continue to struggle and strain for release. Perhaps one should say in such instances that the struggle is a response to the confinement of the stock and experimental situation—it is the 'freedom reflex' of Pavlov (**6**, p. 11). And that the flexion CR has no relation to it whatsoever, but develops, so to say, at an entirely different level. We have even had dogs which did not struggle at first but which developed struggle reactions

at advanced stages of conditioning. In addition to cases like these, there are, of course, those subjects that follow the more orthodox trial-and-error procedure and gradually eliminate their random activity.

Entirely apart from individual differences of this kind, the struggle response is peculiarly dependent upon the intensity of electrical stimulation used for the unconditioned stimulus. Any dog—even one which has never previously struggled in the apparatus—can be made to do so by an increase in the E.M.F. to a point beyond that necessary for obtaining the unconditioned flexion lift of 4 inches required by the experimental procedure. It is not impossible *to produce persistent conditioned struggle in a normally quiet animal* by a few extra-strong stimulations of this sort. It is also possible to increase the violence of struggling in any dog to the point where he may do himself bodily harm. Yelping, defecation, micturition, and biting or chewing at the stock may be added to the behavior-picture under such circumstances.

Some of the ways in which the struggle response can be affected by the conditioned stimulus during the process of training are illustrated in Figs. 1, 2, and 3, which are reproductions from kymographic records made during the conditioning of 3 different animals. The top line, *R*, in all 3 figures is the respiration, *LF*, *RF*, *LR*, and *RR* show the movements of the left-front, right-front, left-rear, and right-rear limbs, and *B*, *S*, and *T* are the buzz, shock, and time lines. In each of the 3 cases pictured the *RR* foot only was shocked.

In Fig. 1 is shown the more common or ordinary reaction to the buzz stimulus, which takes place during the early stages of training. Shortly after the incidence of the buzz, the subject has rapidly lifted all 4 feet in a series of 'pumping' movements. His respiration has been correspondingly disturbed. The extremely rapid breathing between most of the trials indicates that the dog was panting. Here, then, is a typical picture of the conditioned struggle response.

In Fig. 2 the reverse phenomenon is illustrated. Before the beginning of 3 of the 4 trials pictured, the subject was in the act of struggling. The record clearly shows that the

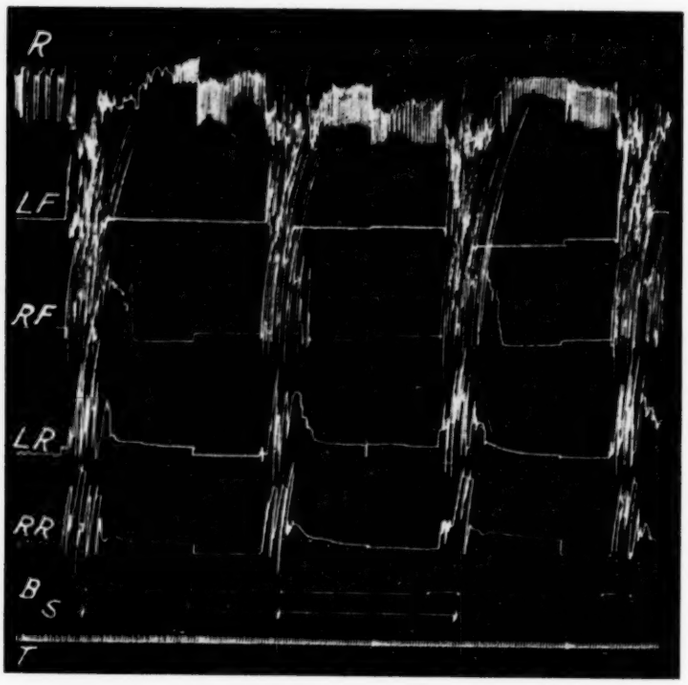

FIG. 1

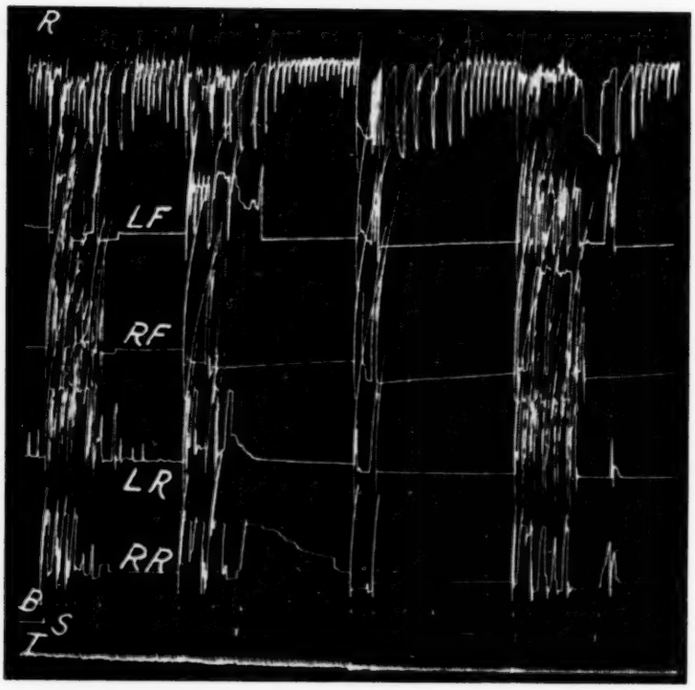

FIG. 2

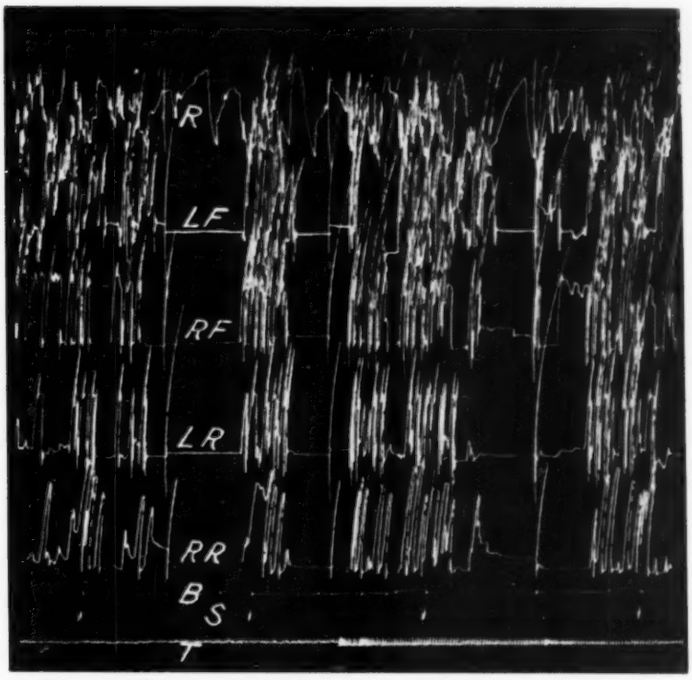

FIG. 3

Figures 1, 2, and 3 show struggle reactions as recorded in original kymographic tracings. Each figure is from the records of a different animal. In all 3 figures, the top line, *R*, is the respiration, *LF*, *RF*, *LR*, and *RR* show movements of the left-front, right-front, left-rear, and right-rear feet, and *B*, *S*, and *T* are the buzz, shock, and time lines.

Figure 1 gives a typical picture of conditioned struggle. Shortly after the incidence of the buzz, the subject has rapidly lifted all 4 feet in a series of 'pumping' movements. His respiration has been correspondingly disturbed. The extremely rapid breathing between most of the trials indicates that the dog was panting.

Figure 2 shows the inhibition of struggle to the buzz stimulus. The phenomenon is just the reverse of that presented in Fig. 1.

Figure 3 illustrates a third sort of response—or the lack of it—to the buzz. The struggle which was well under way before the conditioned stimulus was introduced (in 3 of the 4 trials shown) has continued through and beyond each of these 3 trials. There is some indication, notably in the trial farthest to the right, that the buzz may have served to inhibit or slow up the struggle for a fraction of a second, but the hesitation, if actual, was of short duration.

struggle was either reduced or eliminated entirely by the buzz stimulus. The trials pictured are numbers 283–286 inclusive, in the middle of a longer training series.[4]

Figure 3 illustrates a third sort of reaction—or the lack of it—to the buzz. The struggle which was well under way before the conditioned stimulus was introduced (in 3 of the 4 trials shown) has continued through and beyond each of these 3 trials. There seems to be some indication, notably in the trial farthest to the right, that the buzz may have served to inhibit or slow up the struggle for a fraction of a second, but the hesitation, if real, was of short duration. The reactions pictured in this figure took place on the sixth day of conditioning during trials 116–119 inclusive. Struggle had persisted in this animal since the very beginning, and continued to a point considerably beyond the trials shown.

Since struggling itself is so unpredictable and irregular a phenomenon, as demonstrated both from these figures and from the other evidence presented, there is considerable doubt whether it should be regarded as the raw material from which localized flexion CRs are ultimately refined.

'*Hypotheses*' *in Conditioning.*—Over and above the fact that struggling can be eliminated after the early stages of training, that it can be introduced in previously quiet animals at advanced stages of training, or that it can continue throughout the entire training period, there remains the question of its speed or rate of elimination or onset. That this rate is frequently gradual there can be no doubt. But it can also be rapid, as is shown by the performances of a number of animals from our laboratory which dropped the struggle reaction with astonishing speed. They would struggle regularly up to a certain point in the training, and then abruptly and without warning, they would eliminate the struggle for all subsequent trials. In one or two cases struggle has also been introduced at a late stage of training in the same abrupt manner, while in others it has come and gone.

Figures 4 and 5 are reproductions from kymographic records which illustrate the sudden elimination of struggle.

[4] Beritoff (1, p. 366) reports a 'negative action' of the conditioned stimulus which he regards as the after effect of an earlier unconditioned stimulus.

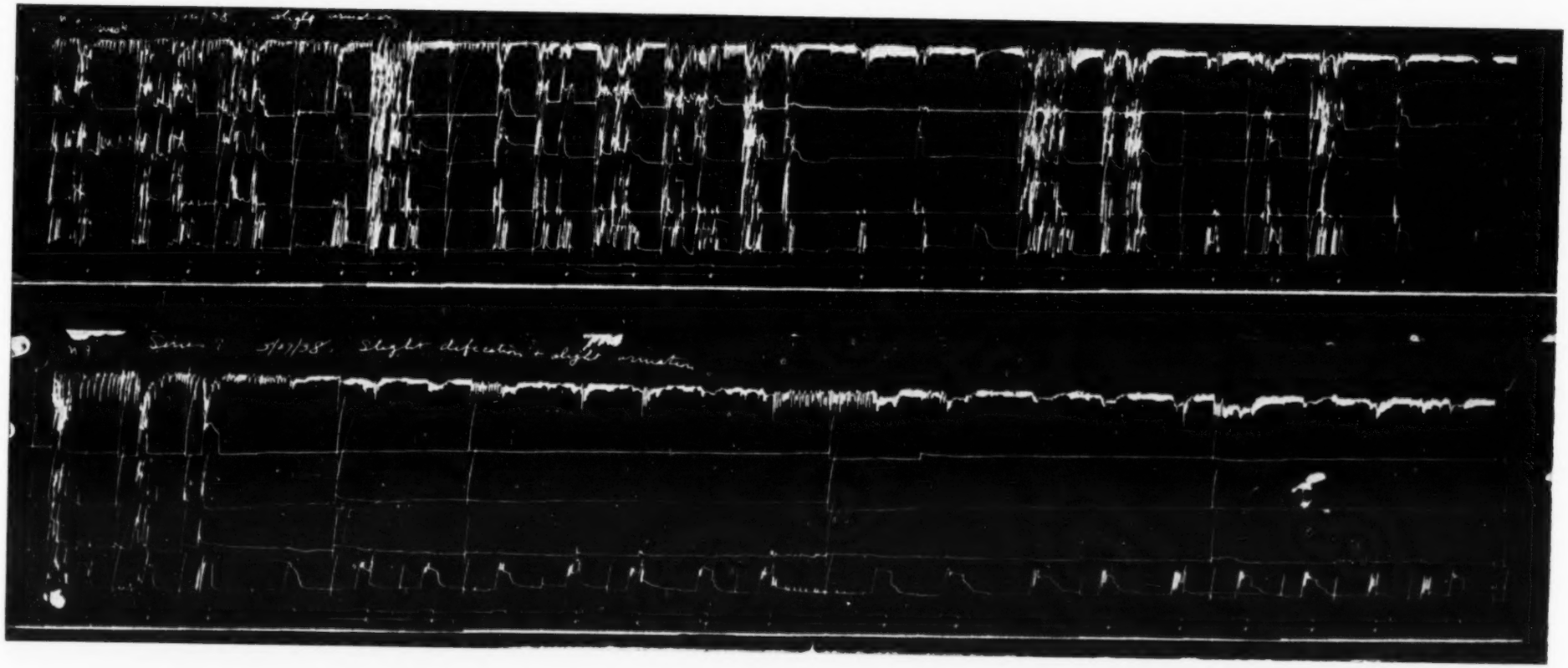

FIG. 4

Figure 4 shows 40 trials (2 days) of training near the middle of a longer training series, which illustrate the sudden elimination of the struggle response. Since the change persisted in this instance (and in other similar cases) it may be considered as an example of 'hypothesis behavior' in conditioning. Before the upper record in Fig. 4 was made, the subject had received exactly 100 conditioning trials. The shift to adaptive flexion conditioning took place in trials 101–120 and—except for the first 3 trials on the following day—was well enough established to remain.

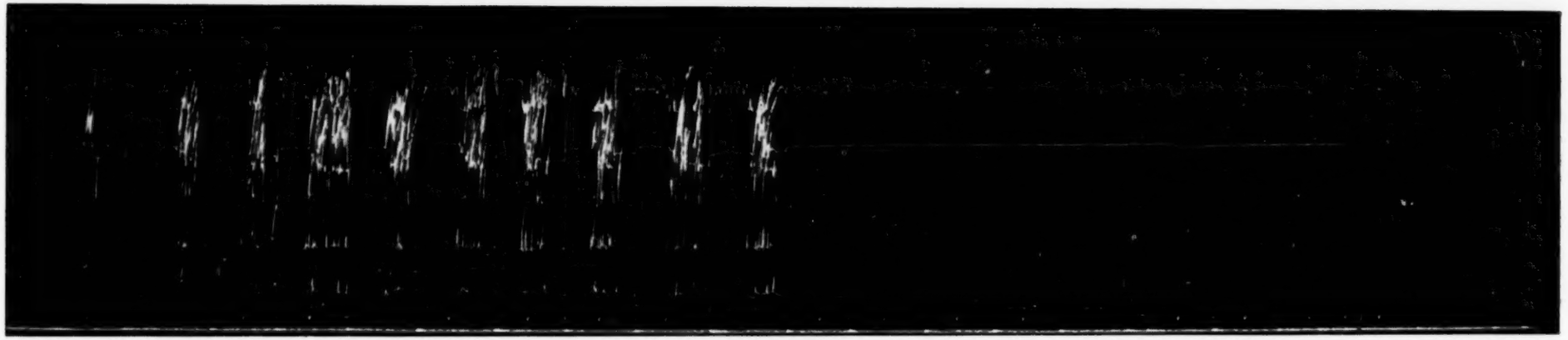

FIG. 5

Figure 5 shows an even more rapid elimination of struggle in a different animal, which occurred on the very first day of conditioning. During 400 subsequent trials (20 days) there was never any return of the struggle response.

Figure 4 depicts 40 trials (2 days) of training near the middle of a twenty-day training series. Before the upper record in Fig. 4 was made, the subject had received exactly 100 conditioning trials. The shift to adaptive flexion conditioning took place in trials 101–120 and—except for the first 3 trials on the following day—was well enough established to remain. Figure 5 shows an even more rapid elimination of struggle in a different animal, which occurred on the very first day of conditioning. During 400 subsequent trials there was never any return of the struggle response.

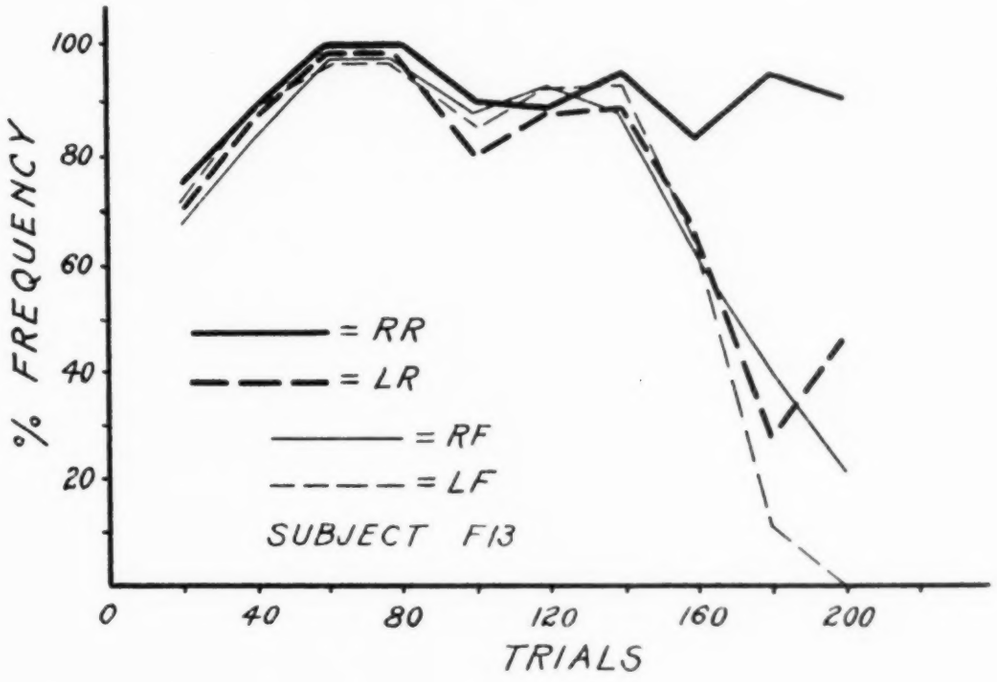

FIG. 6. Here are plotted a series of 4 'learning curves,' one for each foot of the subject. The percent frequency of responsiveness on the ordinate is the percent of CRs in each daily group of 20 stimulations. Although the *RR* foot only was shocked, CRs were counted for any foot providing that foot was lifted at all during the 9/5 sec. of buzz preceding the unconditioned shock stimulus. The sudden dropping out of the CRs from the 3 non-shocked feet after 140 trials, is evidence of the phenomenon of insight, or a change in 'hypotheses.'

A second way of demonstrating the shifting phenomenon is shown in Fig. 6. Here are plotted a series of 4 'learning curves,' one for each foot of subject F13, as that subject made a sudden and 'successful' change from struggle to non-struggle *CR*s. The percent frequency of responsiveness given in this figure is the percent of *CR*s in each daily group of 20 stimula-

tions. Although the *RR* foot only was shocked, *CR*s were counted for any foot providing that foot was lifted at all during the period of the buzz stimulus. The graphs for the *LF*, *RF*, and *LR* feet may hence be considered as 'error' curves since they depict unnecessary or incorrect responses. But the graph for the *RR* foot is an increasing-score curve of efficiency of learning.

It will be noted that the subject responded with a frequency of around 70 percent in all 4 feet during the first group of 20 stimulations. Thereafter the frequency increased to about 100 (*i.e.* the dog struggled) on nearly every trial. After 140 trials had been given, the animal abruptly reduced the responsiveness of the 3 non-shocked feet, although she maintained a level of near 100 percent in the *RR* member. She therefore developed a localized flexion *CR* with all the suddenness characteristic of the phenomenon of insight.

Changes in behavior like these, which are so far-reaching and significant to the organism that they include violent effort from the whole body at one time, and complete removal of the violence immediately afterwards require some explanation other than 'chance,' 'random activity,' or 'trial-and-error.' During the struggle behavior the subjects engage in a kind of behavior which, although definitely circumscribed by the experimental situation, is vastly different from that which follows. Either of these two sorts of responses, or both of them, have been found to persist in different animals throughout long periods of the training. We regard them as sequences of activity surely as regular and as clear-cut as the rat's consistent turning to the right or to the left in the discrimination or choice situation. It seems to the present writers, therefore, that the most satisfactory interpretation is in terms of Krechevsky's 'hypotheses.'

There are other 'milder' sorts of 'hypothesis behavior' which do not involve as widespread, or as violent activity as that included under the heading of struggle. An animal may develop the habit, for example, of regularly raising his left-front foot just before he lifts his shocked (*e.g.* the right-rear) foot. With continued training this particular pattern of

'(left-front)—(right-rear),' may give way to some other combination of movements, in which the *RR* flexion is necessarily a part. Or, it may shift directly to the 'adequate' response of '(right-rear) alone.' Barking, whining, respiratory changes without vocalization, and other reactions not recorded by the present technique, enter into still other sequences of behavior which have been observed. The change from one of these combinations to another is not infrequently sudden and abrupt.

When applied only to cases like that shown in Fig. 6, where the shift in behavior leads the animal to 'the correct response' —that is, from struggle to non-struggle behavior with adaptive flexion CRs in the shocked limb—the change can be regarded as a kind of insight. But if the shift goes in the wrong direction, or if there are several sudden changes, including reversals and repetitions, then the doctrine of hypotheses in learning is the more adequate of the two.

IV. Summary

The commonly accepted view with reference to the flexion conditioning of dogs is that the process begins with the development of conditioned struggle by the animal. As training progresses, the specialized or adaptive flexion CR emerges gradually from out of the generalized conditioning of the early stages. Since struggling in the confining experimental situation is about the most 'random' activity which is permitted the subject, he may be said to develop the localized CR from the raw material of random movements. Such an account is, in effect, an application of the trial-and-error theory to the conditioning situation. Observations of the present writers, based upon the conditioning of nearly 60 animals, have shown that this description of the conditioning process cannot be generally applied, although it does hold in many instances. In a majority of the cases studied, kymographic records were obtained of the movements of all 4 feet and of the respiration during the flexion conditioning of only one foot (or of two feet).

1. Some animals, although they develop the conditioned flexion response, may never struggle at all.

2. Others may struggle almost continuously, probably in response to the confining experimental situation, but develop the conditioned flexion CR as an entirely independent reaction. They may never eliminate struggle even if given hundreds of conditioning trials.

3. Occasionally a subject may develop conditioned struggle *after* he has developed the flexion CR without struggle.

4. Struggling, or 'random behavior' in the conditioning situation, is also a function of the intensity of electrical stimulation, and can be produced or eliminated by adjusting the unconditioned stimulus accordingly.

5. The stopping of struggle or its onset frequently takes place with a suddenness or abruptness which is characteristic of Gestalt insight.

6. Other persisting patterns of activity—not violent enough to be classified as struggling—have also been observed in response to the conditioned stimulus. These may involve peculiar temporal sequences of lifting one or more of the non-shocked feet. They may include vocalizing and other reactions. Sometimes the shift from one such combination of movements to another is also sudden or abrupt.

7. To account for these rapid shifts between different patterns of persisting behavior, the doctrine of 'hypotheses' in learning has been applied.

8. It is not necessary to suppose, therefore, that insight or 'hypotheses behavior' is something which is entirely incompatible with the traditional 'conditioned reflex,' or that the two represent distinct and separate kinds of learning.

(Manuscript received December 18, 1939)

V. References

1. Beritoff, J. S., On the fundamental processes in the cortex of the cerebral hemispheres. II. On the principal cortical elements in the arcs of the individual reflexes, *Brain*, 1924, **47**, 358–376.
2. Kellogg, W. N., Davis, R. C., and Scott, V. B., Refinements in technique for the conditioning of motor reflexes in dogs, *J. exper. Psychol.*, 1939, **24**, 318–331.
3. Krechevsky, I., 'Hypotheses' versus 'chance' in the pre-solution period of sensory discrimination-learning, *Univ. Calif. Pub. in Psychol.*, 1932, **6**, No. 3.
4. Krechevsky, I., 'Hypotheses' in rats, *Psychol. Rev.*, 1932, **39**, 516–532.
5. Krechevsky, I., The genesis of 'hypotheses' in rats, *Univ. Calif. Pub. in Psychol.*, 1932, **6**, No. 4.

6. Pavlov, I. P., *Conditioned reflexes.* (Trans. by G. V. Anrep.) Oxford Univ. Press, 1927, 430 p.
7. Spence, K. W., The nature of discrimination learning in animals, *Psychol. Rev.*, 1936, **43**, 427–449.
8. Tolman, E. C., *Purposive behavior in animals and men*, New York, Century, 1932, 463 p.
9. Tolman, E. C., and Krechevsky, I., Means-end-readiness and hypothesis. A contribution to coomparative psychology, *Psychol. Rev.*, 1933, **40**, 60–70.
10. Witkin, H. A., 'Hypotheses' in rats: a critique, *Psychol. Bull.*, 1938, **35**, 638.
11. Wolf, I. S., and Kellogg, W. N., Changes in general behavior during flexion conditioning and their importance for the learning process, *Amer. J. Psychol.*, 1940, (in press.)

THE RELATIONSHIP BETWEEN PERFORMANCE LEVEL AND BODILY ACTIVITY LEVEL *

BY G. L. FREEMAN

Northwestern University

In a previous report [1] the writer suggested a relationship between the degree of overt 'phasic' performance and the amount of supporting 'tonic' reactivity, in which optimal conditions were alleged to obtain when the excitation level neither falls below or rises above certain limits. Direct alternation of excitation background by experimentally induced muscular tensions has tended to confirm the theoretical analysis.[2] The purpose of the present experiment is to study this same relationship indirectly by encouraging subjects to exert varying degrees of effort in doing a task. That such 'effort' will be reflected in the postural accompaniments of performance is already well recognized.[3] Other studies [4] have shown that the palmar skin resistance may be used as an index of postural change. Specifically then, we are herein concerned with the level of palmar skin resistance when individuals are performing above, below and at their normal 'congenial' pace such tasks as finger oscillation and reaction time.

Experiment I

The first series of observations was made upon twenty undergraduates under two conditions: (1) wherein the individual set his own standard of performance; and (2) wherein

* Studies in Muscular Tension, Series 2, No. 14. The author wishes to acknowledge his indebtedness to Miss R. Buchanan and Mr. M. Schisler, who assisted in carrying out this research.

[1] G. L. Freeman, The postural substrate, *Psychol. Rev.*, 1938, **45**, 324–333.

[2] G. L. Freeman, The optimal muscular tensions for various performances, *Amer. J. Psychol.*, 1938, **51**, 146–151.

[3] J. J. B. Morgan, The overcoming of distraction and other resistances, *Arch. Psychol.*, 1916, No. 35.

[4] G. L. Freeman and R. M. Simpson, The effect of experimentally induced muscular tensions upon palmar skin resistance, *J. gen. Psychol.*, 1938, **18**, 319–326.

the standard of performance was set an arbitrary number of points (1 *P.E.*) higher than the individual's averages for the first condition, with failures in meeting the raised standard called to his attention. Both conditions were preceded by fifteen minutes of relaxation, or until measure of palmar skin resistance showed no marked change for five minutes.[5] The

TABLE I

SHOWING AVERAGE REACTION TIMES (IN SECS.) AND PALMAR SKIN RESISTANCES (1000 OHMS) OF TWENTY SUBJECTS UNDER NORMAL AND SUPERNORMAL 'EFFORT' CONDITIONS

Subject	Normal Effort			Supernormal Effort		
	Resting Resistance Level	Working Resistance Level	Reaction Times	Resting Resistance Level	Working Resistance Level	Reaction Times
Fe	55	37	.125	55	35	.108
Bu	26	22	.104	35	25	.100
Be	65	40	.120	65	42	.117
Sp	25	15	.111	45	25	.118
Br	50	36	.110	55	39	.090
Ba	59	45	.112	67	44	.120
Ho	55	40	.087	35	15	.89
Bk	45	25	.104	80	25	.116
Zi	145	135	.140	147	134	.104
Li	45	39	.128	54	35	.122
La	43	30	.145	50	38	.126
Mo	85	45	.120	87	45	.122
Ei	35	25	.110	30	28	.104
Ke	60	50	.091	65	26	.096
Dr	65	25	.112	65	22	.120
Ga	45	25	.080	45	25	.076
Mc	37	19	.94	45	14	.091
De	49	25	.110	55	35	.88
Ul	90	81	.127	95	85	.114
Ad	30	20	.096	25	15	.096
Ave.	55.45	38.95	.153	59.9	37.6	.185

two conditions were usually presented on consecutive days with ten five-trial periods of auditory reaction tests and ten thirty-second periods of finger oscillation interspersed with thirty-second rest periods. All subjects received an hour's preliminary practice. Instructions for the two experimental conditions were:

a. 'Normal' Competitive Effort.—"There are wide individual differences in speed of (reaction time; finger oscillation)

[5] G. L. Freeman and L. L. Giffin, The measurement of general reactivity under basal conditions, *J. gen. Psychol.*, 1939, 21, 63–72.

and you will naturally want to attain as high a score as possible in comparison with the rest of the class. Scores will be posted.

b. 'Supernormal' Competitive Effort.—"We think you would not be entirely satisfied with your relative standing on the test of yesterday. Before we post scores we are giving another test, wherein we hope you can excell your previous performance. To aid you we have set the standard somewhat above the average you originally attained and will announce

TABLE II

SHOWING AVERAGE FINGER OSCILLATION (PER 30 SEC.) AND PALMAR SKIN RESISTANCE (1000 OHMS) OF TWENTY SUBJECTS UNDER NORMAL AND SUPERNORMAL 'EFFORT' CONDITIONS

Subject	Resting Resistance Level	Working Resistance Level	Oscil.	Resting Resistance Level	Working Resistance Level	Oscil.
Mo	85	30	109.1	87	31	100.4
Bu	26	22	111.4	35	27	137.2
Ei	35	22	125.7	30	26	139.3
Ad	30	21	126.4	25	15	140.5
Bk	45	22	120.9	80	30	122.6
Fe	55	39	106.8	55	25	129.3
Mc	37	19	130.5	45	12	145.9
Li	45	42	104.9	54	23	123.5
Sp	25	20	126.7	45	23	162.2
Br	50	31	101.0	55	30	114.9
Ga	45	26	115.7	45	29	113.4
Ke	60	52	116.2	65	35	156.6
Ho	55	50	144.5	35	15	145.1
De	49	25	92.5	55	31	115.4
Lo	43	35	106.0	50	33	128.8
Ba	59	45	78.2	67	52	97.9
Ul	90	65	103.0	95	60	127.5
Zi	145	140	116.7	147	130	123.8
Be	65	50	83.6	65	34	119.4
Dr	65	28	96.9	65	26	124.7
Ave.	55.45	39.15	110.8	60	34.35	128.4

your success or failure in meeting this standard on each trial made. This will require considerable added effort in which we request your full coöperation."

Results indicate that the second instruction (*b*) was generally successful in motivating the subjects to increased effort or reactivity. As shown in Tables I and II, however, performance does not improve for all subjects under increased 'effort,' though general averages indicate a slight improvement. The resting resistance levels are fairly equivalent for

the two conditions ($r = +.67$), hence it seems possible to compare the degree of reactivity change under different effort incentives. One of the most significant of these comparisons is in relation to the change in performance. If we correlate the amount of performance change (condition *a* over *b*) with the degree of reactivity change exhibited, it appears that for reaction time those who exerted the greatest effort (reactivity) showed the least improvement and often did worse under supernormal conditions ($r = -.41$ for reaction time; $r = -.16$ for finger oscillation).

Various other measures have been applied to the data of this experiment, but due to many factors not under control, have not been found highly significant. In general it seems that persons of high initial resting reactivity (low skin resistance) do not increase performance under raised standards as much as those with a relatively lower initial reactivity level. Under 'supernormal effort' the greater the drop in palmer skin resistance (from initial level) the faster is the time in attaining the supernormal standard ($r = +.56$ for finger oscillation; $+.22$ for reaction time). Also, there is a tendency for persons who showed the greatest drop in skin resistance (resting to working level) to attain the supernormal standard a greater number of times ($r = +.45$ for finger oscillations; $+.17$ for reaction time).

Experiment 2

Because of the many complicating factors which are introduced when the reactivity-effort and performance-quality relationship is studied briefly in a number of individuals, it was decided to make a thorough survey of performance and reactivity on a single person working under all varieties of incentive. The reaction time and palmar skin resistance were recorded simultaneously in a series of one hundred five-trial observations. The observations were made at various times with *S* in various states of alertness; sometimes they were made on rising from bed; again, following an emotional crisis. Naturally no attempt was made to get basal reactivity to comparable levels. Sometimes *S* tried to work hard; some-

times he loafed. He did not examine the records of performance or reactivity during the experiment. In Fig. 1, performance is plotted against reactivity and indicates that for this subject as reactivity increased performance first rose and then fell off. The points marked with numbers 1, 2, etc., indicate observations where *S* made a report on his psychological condition; these reports are summarized on the graph

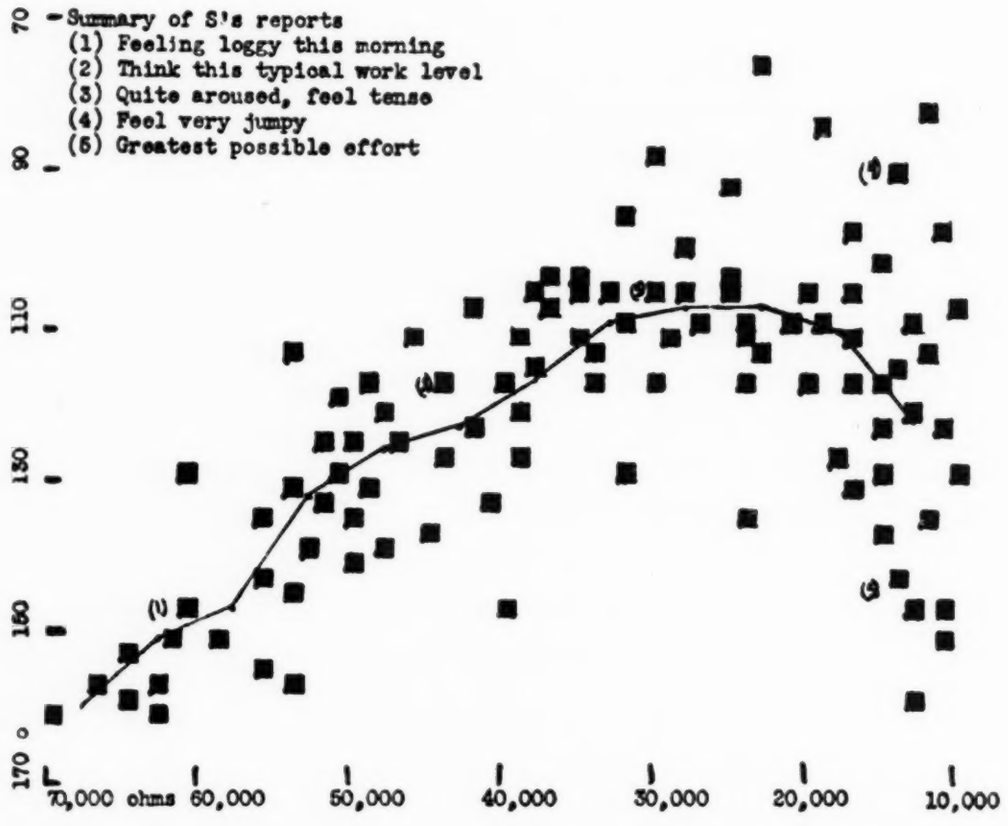

FIG. 1. The relationship between palmar skin resistance and reaction-time performance in a single subject under various states of alertness.

and indicate that *S* was nearly asleep at points along the first part of the curve, and excited or 'rattled' at points along the latter part of the curve.

EXPERIMENT 3

The final experiment was concerned with increasing and decreasing motivational effort in reference to the subject's 'normal' or 'congenial' pace. Measures of skin resistance and reaction time or finger oscillation performance were taken

under conditions (*a*) where *S* was instructed to assume a comfortable working pace and when he was urged (*b*) to work above or (*c*) to work below this congenial pace. Fifteen *S*'s were used, with each of the three work periods (50 reaction time trials; 10 half-minute finger oscillation trials) preceded by a relaxation control.

In general, the results of this experiment indicate that 'effort' to work above the congenial pace is accompanied by increased performance and decreased palmar skin resistance; conversely, the instruction to work below the congenial pace tends to produce lowered performance and increased skin resistance. These effects are more pronounced for finger oscillation than for reaction time performance. On the reac-

TABLE III

SHOWING THE RELIABILITY BETWEEN THE MEANS (CRITICAL RATIOS) OF THE AVERAGE RESISTANCES OF THE REACTION TIME EXPERIMENT AND THOSE OF THE FINGER OSCILLATION EXPERIMENT

	Relaxation and Control	Control and Raised Level	Control and Lowered Level	Raised and Lowered Level
Reaction time.....	1.11	1.07	0.16	1.29
Finger Oscillation.....	1.64	2.56	1.18	2.83

tion time tests, instruction to work slow was sometimes reported to involve more subjective effort than working at the congenial pace, and this is apparently reflected by the drop rather than the expected rise in skin resistance, over the 'normal' control. Such facts, together with a relative lack of success in establishing comparable basal levels of skin resistance in the relaxation period preceding each condition ($r = .47 \pm .11$), makes for many reversals in the general negative relationship between palmar skin resistance and work output. Correlations show, however, that persons of low performance on the finger oscillation test were also low on the reaction time test, and tended to maintain their relative position under the three incentive conditions; the correspondence in relative changes was least exact at the 'normal' pace ($r = .28 \pm .16$), more so at the decreased pace ($r = .45 \pm .14$), and most exact at the increased pace ($r = .62 \pm .11$). The

correlation between average work output and average accompanying skin resistance was higher for finger oscillation (normal pace, .30±.21; raised pace, .48±.14; lowered pace, .49±.18) than for reaction time (normal pace .09±.18; raised pace, .06±.18; lowered pace, .14±.17), but in neither case was the evidence very significant. That the incentive conditions were not successful in raising or lowering the palmar skin resistance significantly is shown by Table III.

Discussion

From the three experiments reported above, and especially for the last one, it is suggested that changes in skin resistance are more related to subjective effort than to performance level *per se*. The fact that it is often difficult for an individual starting from a relatively high level of reactivity to attain a slow working pace, makes for many reversals in the assumed decrease in reactivity and under lowered performance. Furthermore, measures of work-output alone are insufficiently to show the assumed relationship; for while an individual working at high or low excitation levels may turn out the same amount of work as at a more congenial pace, a decrement in the quality of performance can have occurred. Another complicating factor is the possibility of change in work methods at different production rates, which would tend to mask the relationship between effectiveness of performance and reactivity. Later research may conceivably show that the extremes of this relationship are attained only when resting excitation level, from which increments obtain during work, is abnormally high or low. Certainly this is the implication from our second experiment, done on a single subject.

(Manuscript received November 23, 1939)

TRIAL AND ERROR LEARNING IN PARAMECIUM *

BY J. W. FRENCH

Princeton University

There is increasing evidence that the lower animals are not entirely reflex machines. Even the one-celled animals have been shown to modify their behavior as a result of previous activity. Jennings (**6**, Chapt. X) regarded this modification as a change in the physiological state of the organism. He believed that such a change explains why stentor when first subjected to carmine grains will wave back and forth, but upon continued stimulation by the grains will contract or swim away.

Learning in unicellular animals has been studied most extensively in paramecium. Smith (**7**) and Day and Bentley (**4**) attempted to show that paramecia can learn to double on themselves in order to turn around in a capillary tube so narrow as to restrict their free movement. The average time and number of unsuccessful attempts to turn was greatly reduced with practice. They showed that this apparent learning was not caused by an accumulation of CO_2, since the effects of practice were retained after the animals had been swimming in an open dish for twenty minutes. Buytendijk (**3**), however, suggested that this is not true learning but merely a change in physiological state resulting in greater flexibility of the animal. He found that an animal treated in chloroform becomes flexible enough to turn easily in the tube.

That paramecia can learn to associate the avoidance reaction to heat with the stimuli of mechanical shock or light was claimed by Bramstedt (**2**). He stated also that paramecia adapt to the shape of a container, for, on removal from the container, they continue to swim in a course similar to its outline. Bramstedt's observations on the association of light

* From the psychological laboratory, Princeton University.

with heat were confirmed by Alverdes (1). Grabowski (5) on the other hand disagreed with the conclusion that learning is involved and gave evidence that the animals were merely reacting to a chemical change in the liquid arising from the action of heat. Grabowski also believed that the observations on adaptation to the shape of a container are not sufficiently conclusive to warrant the assumption of learning.

Soest (8) was unable to form an association between dark or light with cold. Electric shocks could not be associated with dark, but did produce an avoidance of light. Since the electrodes were remote from the experimental area, and the current flowed through light and dark regions alike, a chemical explanation seems inadequate.

Tschakhotine (9) proved that paramecia will avoid a beam of ultra-violet light, and will continue to avoid the same spot for a half-hour after the beam has been removed. Grabowski's objection that a chemical change occurs in the liquid may be pertinent here.

It is the purpose of the present paper to show that paramecia learn by trial and error. When a paramecium is enclosed in a glass tube large enough to permit free swimming, it avoids contact with the hard smooth surface of the tube. The animal beats back and forth and eventually swims up or down. By swimming down it is able to escape from the open end. The object of the experiment was to observe any change with practice in the time taken to escape. In contrast to Day and Bentley's experiment the flexibility of the animal is not important here, since the tube's diameter was several times the animal's length and permitted easy turning.

Method

The tube used to confine the animals was .6 mm in internal diameter and 4.6 mm in length. Preliminary observations showed that when wider and shorter tubes were used the elements of chance escape appeared to obscure the learning. Since the average length of the animals was only about .2 mm, they could swim about freely within the tube.

A single animal in liquid from his own culture was placed on a piece of glass one inch square. The glass tube, supported vertically in a metal frame, was then lowered over it, so that it was sucked up into the tube by capillary attraction. Since the tube was held slightly above the plate, the animal could escape by swimming out the lower

end. Each of twenty animals was given thirty trials spaced at 15″ intervals. After each trial the liquid in the tube was removed, so that no chemical could accumulate during the practice period.

Except for a few isolated trials the animals seemed well motivated, for their activity within the tube was vigorous. The room was maintained at 26° C., since the avoidance of contact is stronger in warm than in cold water (**6**, p. 99).

Results

The results are given in Fig. 1, which shows the median score of the group of twenty paramecia for each trial, and also

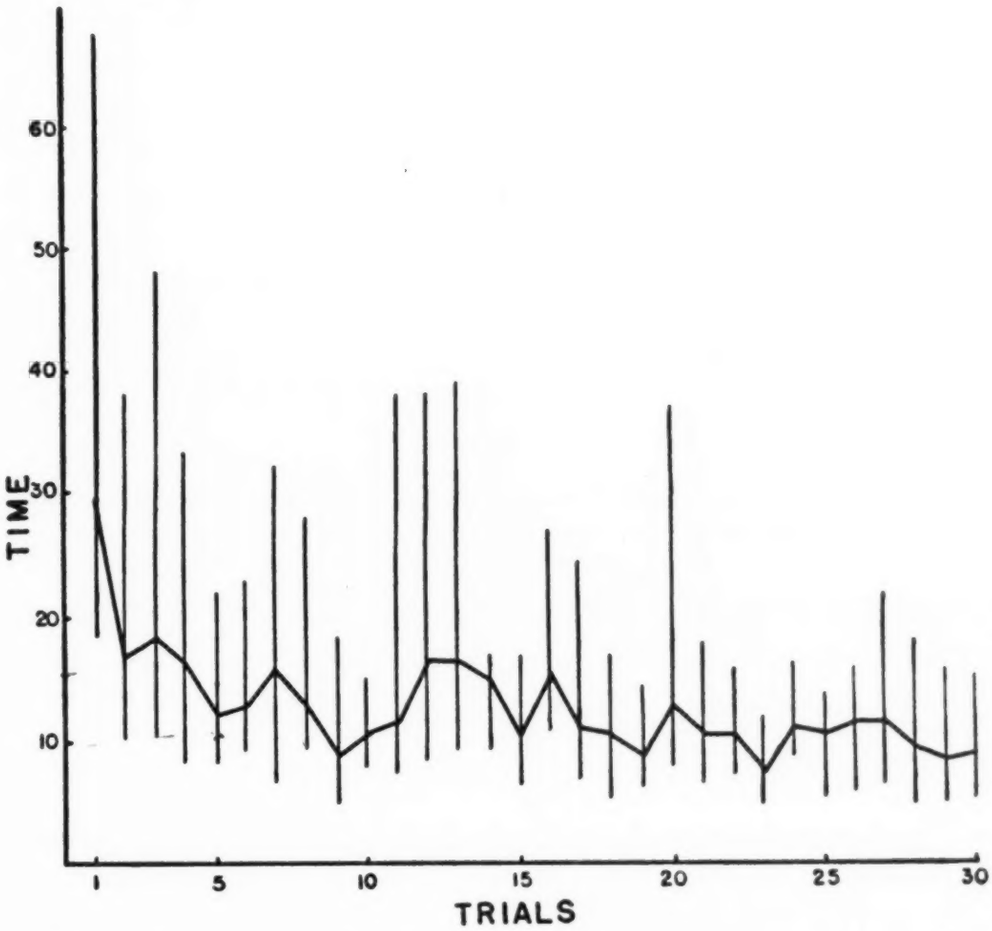

Fig. 1. Learning curve for the group of twenty paramecia. The median time for escape from a glass tube is given in seconds. The ends of the vertical lines mark the upper and lower quartiles of the group for each of the thirty trials.

the upper and lower quartiles. In general the curve tends downward, but there is considerable variability. Examination of the individual results revealed consistent improvement during the early trials for ten animals, but for the other ten a decrease in time to escape was doubtful or absent. Two

typical individual curves are given in Fig. 2. For animal no. 1 there is a fairly regular improvement during the first 13 trials with little change thereafter. For animal no. 7 there is merely great variability from trial to trial. Despite the relatively great variability between individuals and from trial to trial of some individuals, the group as a whole shows improve-

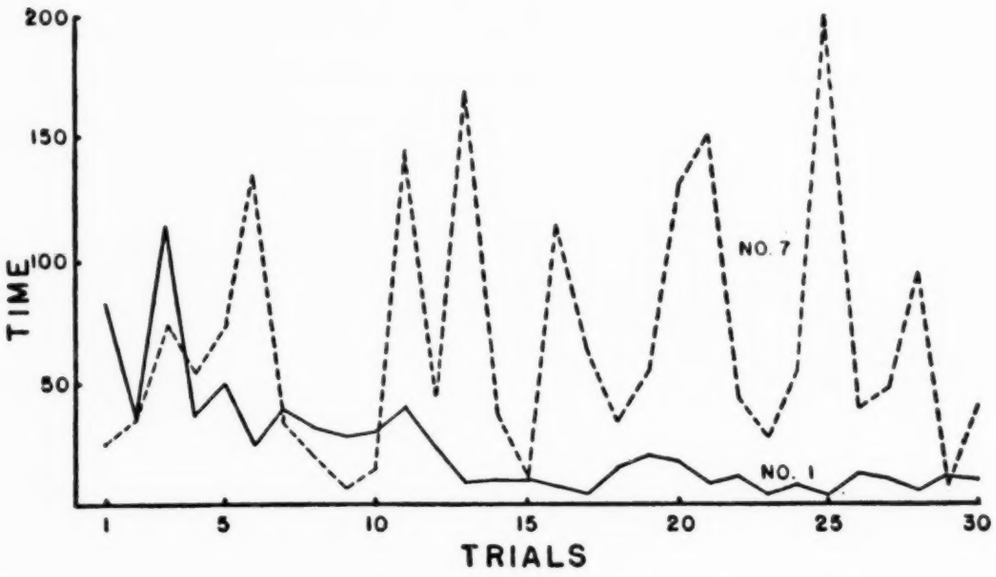

FIG. 2. Learning curves of animals nos. 1 and 7. The time for escape from the tube on each of the thirty trials is given in seconds.

ment. The difference between the mean scores of the group on the first and last trials is statistically significant, being 3.43 times its standard error.

DISCUSSION

The increased speed with which the animals escaped from the tube might result from an increase in general activity. To check on this possibility an independent measure of swimming speed was made before and after the practice period. Each animal was placed over a wire screen having ten squares to the inch, and the number of squares over which it passed during one minute was counted. The average score before practice was 42 and following practice was 41.2. Thus there was no general change in activity. Furthermore, inspection of the activity scores gave no indication of a correlation between activity and efficiency in the learning situation.

Qualitative observations of the performance of the paramecia during the practice period suggested that the improvement is due to learning. During the first trials the animals would alternately beat back and forth in the tube at one level and swim short distances up or down until escape was effected. On the later trials some paramecia were observed to beat back and forth only a few times and then take one long dive to the bottom. Such directed behavior, occurring only in certain animals, is difficult to understand in any terms other than those of learning.

Conclusions

When paramecia are repeatedly enclosed in a glass tube which permits free turning, the average time for escape decreases on successive trials. The change in behavior is not related to general activity and seems best understood in terms of learning.

(Manuscript received December 4, 1939)

References

1. Alverdes, F., Das Lernvermögen der einzelligen Tiere, *Z. Tierpsychol.*, 1937, **1**, 35–38.
2. Bramstedt, F., Dressurversuche mit Paramecium caudatum und Stylonchia mytilus, *Z. vergl. Physiol.*, 1935, **22**, 490–516.
3. Buytendijk, F. J., Acquisition d'habitudes par des êtres unicellulaires, *Arch. néerl. Physiol.*, 1919, **3**, 455–468.
4. Day, L. M., and Bentley, M., A note on learning in paramecium, *J. anim. Behav.*, 1911, **1**, 67–73.
5. Grabowski, U., Experimentelle Untersuchungen über das angebliche Lernvermögen von Paramaecium, *Z. Tierpsychol.*, 1939, **2**, 265–282.
6. Jennings, H. S., *Behavior of the lower organisms*, New York: Columbia Univ. Press, 1906.
7. Smith, S., The limits of educability in paramecium, *J. comp. Neurol.*, 1908, **18**, 499–510.
8. Soest, H., Dressurversuche mit Ciliaten und Rhabdocoelen turbellarien, *Z. vergl. Physiol.*, 1937, **24**, 720–748.
9. Tschakhotine, S., Réactions 'conditionées' par microponction ultraviolette dans le comportement d'une cellule isolée (Paramecium caudatum), *Arch. Inst. Prophyl.*, Paris, 1938, **10**, 119–131.

THE VARIABILITY OF EXTINCTION SCORES IN 'SKINNER-BOX' EXPERIMENTS

BY LLOYD G. HUMPHREYS[1]

Department of Psychology, Northwestern University

An outstanding characteristic of quantitative work using the technique developed by Skinner (**5**), *e.g.*, that of Youtz (**7**), Ellson (**1**), Williams (**6**), and Sackett (**4**), is the wide range of variability obtained. If a means were discovered of decreasing significantly this variability, a considerable saving would be effected in the experimenter's time and in animals required. The purpose of this paper is to discuss two such methods.

Apparatus and Procedure

The apparatus was a duplicate of that used by Ellson (**1**) and Williams (**6**). It consisted, essentially, of four partially sound-proofed boxes into which small, horizontal bars could be introduced, these bars activating food-release mechanisms. Responses were recorded on a constant speed polygraph.

Subjects were 80 male albino rats within the age range of 60 to 100 days. They were obtained from the same source as those of Ellson and Williams.

Habituation to the boxes and to the sounds of the apparatus, and acquisition of the horizontal-bar habit were carried out with no significant departures, as far as present purposes are concerned,[2] from the routine established by the preceding investigators. Two important changes were made, however, in extinction procedure: whereas Ellson and Williams had used a criterion of five minutes in which no responses were made to terminate extinction and had first extinguished the habit on the day following acquisition, present procedure called for constant 10-minute extinction periods for all rats regardless of their rate of response, the first such period to take place immediately following acquisition. After this first, immediate extinction the rats were placed in their living cages for a half-hour rest and then returned to the experimental boxes for a second learning period. The following day a second 10-minute extinction period was given. It is thus possible to present results for constant extinction periods immediately following learning and the day following re-learning. A comparison of these results with those of Ellson and of Williams should make possible an evaluation of the effects of the two changes in procedure.

---

[1] This research was carried out in the Institute of Human Relations, Yale University. The writer is indebted to the National Research Council for fellowship funds and to Professor Clark L. Hull who directed the work.

[2] A detailed statement of procedure is given elsewhere (**2**).

Results and Discussion

Table 1 presents the means and standard deviations of responses in first and second extinctions for four groups of subjects from each of three separate investigations. The first column contains the present results which were obtained in two 10-minute extinction periods, the first of which came immediately after acquisition; the second and third columns furnish data from Ellson (1) and Williams (6) which were obtained in extinction to a criterion, first extinction being twenty-four hours after acquisition.

TABLE 1

Comparative Data on the Variability of Number of Extinction Responses

Standard deviations both for the Ellson (1) and for the Williams (6) data were computed from their published standard errors of means. Ellson's 'first extinction' groups had previously undergone extinction of a vertical-bar habit.

Groups	Humphreys Data			Ellson Data			Williams Data		
	N	Mean	Sigma	*N*	Mean	Sigma	*N*	Mean	Sigma
First Extinction									
1	20	21.6	10.1	25	23.9	26.9	35	19.0	25.5
2	20	27.7	10.4	25	31.2	19.1	35	35.0	43.8
3	20	31.0	10.5	25	39.6	17.6	35	48.3	26.0
4	20	46.6	9.5	25	43.9	25.5	35	62.7	37.9
Second Extinction									
1	20	17.9	13.0	25	7.6	8.3	33	6.7	8.0
2	20	25.3	14.1	25	14.4	8.8	34	10.4	12.2
3	20	26.6	13.1	25	19.5	19.1	33	12.3	9.8
4	20	33.6	11.1	25	24.4	22.0	34	19.8	16.3

It is evident that the two changes in procedure materially reduced the variability of the extinction results. A mean of 46.6 with a standard deviation of 9.5 can be compared, for example, to means of 43.9 and 48.3 whose standard deviations are 25.5 and 26.0; or means of 31.0 and 33.6 with standard deviations of 10.5 and 13.1 can be compared to those of 31.2 and 35.0 with standard deviations of 19.1 and 43.8. There is, in fact, no case where comparable means are found in which the difference between standard deviations is not in the same

direction. This is true though Ellson, before differentiation of his experimental groups, discarded fifteen rats that gave more than 125 responses in first extinction on either the horizontal- or the vertical-bar.

The difference in variability between immediate and delayed extinction (first and second extinctions, column 4) is relatively small. This indicates that the 10-minute extinction period is the chief factor in producing less variable results. The fact, however, that the differences are consistently in the direction of increasing variability from first to second extinction, in conjunction with smaller means and a tendency toward smaller differences between means, would seem to make advisable the use of immediate extinction as well as the 10-minute period.

The use of a standard period of extinction in place of a criterion may be criticized on the basis that the two measure different things. It can be pointed out, for example, that number of responses over a constant period of time is equivalent to rate, but that Youtz (**7**), Ellson (**1**), and Williams (**6**) found that their rate measure did not differentiate between groups. This is, however, an artifact of their computation of rate. Time to the criterion and number of responses to the criterion are undoubtedly highly correlated, perhaps as highly as their reliabilities allow, yet these experimenters divided the latter by the former to obtain rate. Ratios whose numerators and denominators are highly correlated will differ from each other, in the main, only as the relationship between numerator and denominator departs from linearity. It has elsewhere been shown, moreover, that *rate for a constant period of time* does differentiate between groups and does it more efficiently than number of responses to the criterion (**2**).

Restriction of extinction to a constant period of time in the Skinner-type experiment is equivalent to the use of a constant number of trials in the traditional (Pavlovian) conditioning situation. In eyelid conditioning, for example, the latter is now an almost standard procedure. Here, also, correlations between number of trials to a criterion and frequency of responses for a given number of trials have been reported. A

value of −.87, uncorrected for attenuation, is representative of these (3).

In the light of present findings it is highly probable that the use of criteria could be discarded with profit in other learning experiments. Number of errorless 'runs' in the given number of trials replaces the criterion score, while time and error scores per trial are still available. An advantage of this change, in addition to the promise of more homogeneous results, is that a Vincent procedure is unnecessary in constructing performance curves. Criteria are still required, of course, where the acquisition and break-down of the response are subsidiary to other purposes, such as the determination of thresholds.

Summary and Conclusions

In order to reduce the variability of extinction scores in a 'Skinner-box' experiment two innovations in procedure were made: (1) extinction was limited to constant 10-minute periods, (2) one of which immediately followed acquisition. This is in contrast to the usual practice of terminating extinction by a criterion interval in which no responses occur, *e.g.*, a 5-minute interval, and of first extinguishing the response on the day following acquisition. In comparison with published results obtained in the latter, customary way, the two changes in procedure markedly reduce extinction variability (standard deviations one-half to one-fourth as large). The comparisons further indicate that the 10-minute extinction period is the more important of the two factors, but that immediate extinction also contributes to more homogeneous results. Since rate for a constant interval is probably as valid an indicator of strength of response as number of responses to a criterion, the lower variability makes the change to the former definitely advisable. This suggests that the use of criteria could be profitably discarded in other learning situations as well.

(Manuscript received November 24, 1939)

References

1. Ellson, D. G., Quantitative studies of the interaction of simple habits. I. Recovery from specific and generalized effects of extinction, *J. exper. Psychol.*, 1938, **23**, 339–358.
2. Humphreys, L. G., The strength of a Thorndikian response as a function of the number of practice trials (to appear).
3. Humphreys, L. G., Measures of strength of conditioned responses (to appear).
4. Sackett, R. S., The effect of strength of drive at the time of extinction upon resistance to extinction in rats, *J. comp. Psychol.*, 1939, **27**, 411–432.
5. Skinner, B. F., *The behavior of organisms*, New York: D. Appleton-Century Co., 1938.
6. Williams, S. B., Resistance to extinction as a function of the number of reinforcements, *J. exper. Psychol.*, 1938, **23**, 506–522.
7. Youtz, R. E. P., Reinforcement, extinction, and spontaneous recovery in a non-Pavlovian reaction, *J. exper. Psychol.*, 1938, **22**, 305–318.

A STUDY OF CHORD PREFERENCE IN A GROUP OF NEGRO COLLEGE WOMEN

BY ORAN W. EAGLESON AND LILLIAN E. TAYLOR

Spelman College

Purpose

Previous studies have suggested that there is probably no superiority of any single tone from the standpoint of æsthetic appreciation, while there is a rather marked preference for certain combinations of tones. The work of Major (1) offers data in support of the first statement, and the studies of Valentine (2) and others propose evidence for the latter.

This study was undertaken to obtain data on the preference with respect to pleasantness that a group of Negro college women revealed for twelve different tonal combinations. Since Valentine's study presents data for a group of eighty-four White college women, it was also possible to make a racial comparison.

Procedure

Seventy-five Negro college women mainly of sophomore standing were the subjects in this study. Data were collected from about twenty-five subjects at a time.

The following chords were presented on the piano by the method of paired comparisons: minor second, major second, minor third, major third, fourth, fifth, minor sixth, major sixth, minor seventh, major seventh, octave and triad. These chords were constructed within the octave g° to g′ in the same way Valentine set up the intervals for his study; however, in the latter's study the tritone was used, but in the present study this interval was replaced by the triad. The pattern that was followed in constructing the chords may be seen in the following examples: minor second—c′d♭; major second—c′d′; minor third—b°d′. In other words, the pattern was that of alternating between increasing the upper tone or decreasing the lower one a semitone. Each interval was compared twice with each of the other eleven, thus making 132 comparisons in all. It should also be pointed out that Valentine employed the method of single stimuli in presenting his intervals for judgment.

Results

The following table has been prepared to show the preferences of the subjects of both racial groups.

Rank	Data for 75 Negro Women by Method of Paired Comparisons	Data for 84 White Women by Method of Single Stimuli
1	Triad	Major Third
2	Minor Third	Minor Third
3	Minor Sixth	Major Sixth
4	Major Sixth	Octave
5	Fifth	Minor Sixth
6	Fourth	Fourth
7	Major Third	Tritone
8	Octave	Fifth
9	Major Second	Major Second
10	Minor Seventh	Minor Seventh
11	Major Seventh	Major Seventh
12	Minor Second	Minor Second

It is not surprising that the Negroes preferred the triad above all others, for this is consistent with previous work in which the triad has been employed. In comparing the minor with the major chords there seems to be little difference as to which are preferred; however, the slight difference that does exist is in favor of the minor chords.

In comparing the two racial groups one may see from the table that there is more agreement concerning the chords thought not to be so pleasing than for the ones considered most pleasant. In fact, there was perfect agreement for the last four ranks namely, 9, 10, 11, and 12. Excluding the triad and tritone from the data, the rank order correlation was determined for the other seven chords selected for the first eight rankings by the two racial groups. As one may see, these chords are the following: minor third, minor sixth, major sixth, fifth, fourth, major third and octave. The correlation for the two racial groups' ranking of these seven chords is +.04 which indicates very little agreement on the judgment of these chords. It is interesting to note that the major third was the most pleasing to the White subjects, but was ranked seventh by the Negroes.

In considering the racial comparison one must keep in mind the differences in the procedures of the two studies.

Summary and Conclusions

1. Seventy-five Negro college women were requested to judge by the method of paired comparisons the pleasantness

of twelve chords played on the piano. These data were compared with data furnished by Valentine from White college women. The data for the latter study were obtained by the method of single stimuli.

2. The Negro subjects preferred the following chords presented in order from the most to the least pleasing: triad, minor third, minor sixth, major sixth, fifth, fourth, major third, octave, major second, minor seventh, major seventh and minor second.

3. Judgments of the White and Negro subjects were identical as to the four chords that should be ranked least pleasing, but there was little agreement between the two racial groups as to which chords were the most pleasing. The chords on which there was perfect agreement were the major second, minor seventh, major seventh and minor second.

(Manuscript received November 24, 1939)

References

1. Major, D. R., On the affective tone of simple sense impressions, *Amer. J. Psychol.*, 1895, 7, 57–77.
2. Valentine, C. W., The æsthetic appreciation of musical intervals among school children and adults, *Brit. J. Psychol.*, 1913, 6, 190–216.

DISCUSSION

'CENTRAL' VS. 'PERIPHERAL' LOCUS OF SET; A CRITIQUE OF THE MOWRER, RAYMAN, AND BLISS 'DEMONSTRATION'[1]

BY G. L. FREEMAN

Northwestern University

The phenomena of 'set,' long recognized by introspectionists, are belatedly coming to the attention of psychologists who use more 'objective' methods of measurement. This has resulted in a fruitful line of research, the data of which do not seem to have been properly assimilated by certain workers in the field. A recent instance of this is found in Mowrer, Rayman, and Bliss's (10) alleged demonstration of a 'central' locus of set. To me, their report not only fails in its announced purpose, but also threatens to involve a very promising field of research in a futile 'priority' controversy such as has long raged concerning the chicken and the egg. This I would discourage at all costs.

As a person who has found differentiated motor accompaniments for various sets, I might be expected to champion the all-peripheral view of set, but such is not the case. Is any psychological phenomenon all sensory or all motor, all central or all peripheral? I not only doubt it, but am so thoroughly convinced the majority of psychologists agree on this point that I will not waste space in citing the experimental and theoretical reasons which support this position. Just as Dewey showed the impossibility of parting stimulus from

[1] From the Laboratory of Psychophysiology, Northwestern University. The writer wishes to express his appreciation to Dr. L. Harold Sharp for constructive suggestions in the preparation of the manuscript, and to Dr. Lloyd G. Humphreys for a critical reading of it.

This manuscript has been submitted to O. H. Mowrer who indicates that he has further experimental work in progress and that he would prefer to wait until this work is completed before he attempts to meet the criticisms raised in the following paper.—EDITOR.

response, and just as Feré showed the inseparability of sensation and its dynamotor concomitants, similarly we are persuing a scientific will-o'-the-wisp in attempting to establish the priority of central or peripheral determinants of set. Both central 'neural' conditions and peripheral 'muscular' conditions are involved in any set, and if no apparent muscular reactions are found to accompany a set, this no more proves that the locus of set is exclusively central than does the presence of such motor reactions prove that its locus is exclusively peripheral.

But more than a logical criticism can be made of the alleged demonstration of the central locus of set by Mowrer, Rayman, and Bliss. These workers claim to have controlled the peripheral or motor factors by instructing subjects to make the same overt response (lifting a finger) to two different external stimuli. They then varied the number and order of presentation of these stimuli so as to build up different expectancy sets, assuming that since finger preparation (motor set) was the same for both expectancies, lengthened reaction time due to shift of stimulus would indicate the independent variation of 'central' determinants of set. The results of one of their experiments support their hypothetical predictions, while those of another do not. That is, when a sequence of identical stimuli was sufficiently long to shift expectancy in favor of preparation for that type of stimulus alone, the change to presentation of the other stimulus resulted in a lengthened reaction time; but, when alternate presentations of the two different stimuli were followed by an unexpected second presentation of the immediately preceding one, reaction time was shortened rather than lengthened. Mowrer, Rayman, and Bliss are unable to account for this apparent contradiction in terms of their 'central' set hypothesis, but state that "fatigue, confusion, loss of interest, or some other influence, was operative." In the following paragraphs I shall attempt to give a more adequate interpretation of their results.

Mowrer, Rayman, and Bliss admit that their experiments would have permitted differential motor adjustments of the respective sense organs, but they airily dismiss the possibility

of their occurrence as "very remote . . . since movement of the ears could not be made similarly (as might variations in visual sense organ adjustment) to account for the increased latency of the reactions to the first of a series of tones following a series of lights." It is precisely because of this unchecked assumption that their 'demonstration' of the set-locus breaks down. The lengthened reaction time to first presentations of tone stimuli may have been due to either the unprepared state of the *tensor tympani* or to the absence of general motor adjustments favoring hearing. In any event, the motor adjustment made by the eyes in anticipating visual stimulation presumptively creates a cortical gradient which is less adequate for rapid auditory reaction, and which can be changed only by removal of the visual-motor preparation. In my opinion, Mowrer, Rayman, and Bliss's ruling out of motor factors as component determinants of set, by the method of proclamation rather than by adequate control, leaves us with no alternative explanation other than the soul hypothesis masquerading in intra-neural false whiskers.

How is it that allegedly different central (neural) expectancies for sound and light are built up and maintained? We would agree that the responsible factors are neither the covert set of the muscles to be involved in the overt reaction—since finger preparation *per se* is the same for both stimuli, nor the neural traces of the different stimuli—since these, in and of themselves, are relatively short-lived. Outlasting postural adjustments involved in seeing and hearing, however, provide the most probable basis for differentiated central 'expectancies,' since backlash proprioception from these adjustments, operating after the manner of Bok's reflex circle, could establish a cortical dominant of lowered threshold for an about-to-be-presented stimulus. To my way of thinking, it is precisely because serial presentation builds up motor (sense organ) adjustments specific to one type of stimulus that the unexpected occurrence of the other stimulus causes a lengthening in reaction time. The subject is set to see rather than to hear, and the unprepared sense organ must be appropriately adjusted before a differential backlash effect on cortical centers can

structure a central 'expectancy' and so assure the most rapid response. When conditions favor equally the occurrence of both light and sound stimuli at a rate too rapid for correspondingly appropriate sense organ adjustments, the subject presumably attempts to maintain both types of preparation at equivalent levels. An unexpected re-presentation of the sound stimulus, as in Mowrer, Rayman, and Bliss's alternation experiment, now results in a faster rather than a slower response. The so-called 'negative' character of this result presents no difficulty in terms of our hypothesis since the shortened reaction time to a second presentation of the same stimulus is a direct consequent of optimally localized proprioceptive residuals [2] (or after-contractions) from the immediately preceding sense organ adjustment, these acting automatically to increase the excitation level of the corresponding central expectancy. Certainly, there need be no differential pattern of antecedent tension in the reacting finger for different expectancies, but there will be a general increase in tension when the subject is set, and the backlash excitation resulting therefrom will be canalized centrally to reinforce the dominant developed by weaker but differentiated reflex sense organ adjustments.

Whatever may be said of the relative merit of my explanation of the Mowrer, Rayman, and Bliss results, it should be obvious that their experiments are by no means a crucial demonstration of the central locus of set. Their demonstration is crucial only if it is granted that the motor components of their induced sets were localized exclusively in the reacting finger, and that only a quantitative difference existed between these covert sets and the overt response. By assuming that antecedent motor adjustments are qualitatively like those revealed in the overt response of finger lifting, these authors

[2] It has already been shown (11) that increased output results when task resumption coincides with the specifically localized maximal tensions residual from the previous work, whereas decrement occurs under less optimal interval conditions. The reflex after-contractions of sense organ reaction act in a similar manner, presumably, so that if the interval between alternate stimulations is either too short or too long to permit coincidence of optimal proprioceptive residuals with appropriate exteroceptive excitation, longer rather than shorter reaction time would be recorded for re-presentation of the same stimulus.

have failed to control or check upon probable concurrent sense organ adjustments and more widespread postural tensions having differential effects. In earlier reports (**3**, **4**) I have indicated that the overt (phasic) reaction often differs from its supporting (tonic) set qualitatively as well as quantitatively, and that an overt reaction focus can develop upon different preparatory backgrounds. A crucial experiment for demonstrating the central locus of set would be one which showed (by action potentials, eye movement records, or otherwise) that the incipient motor adjustments forming the non-focal parts of the tonic preparatory background did not shift appropriately during the building up of different stimulus expectancies.[3]

A further exception must be taken to the report in question. In a section on 'related experiments,' its authors cite the fact that Hathaway (**8**) was unable to detect any electrical changes in the muscles of the arm which coincide with the development of a readiness for simple reaction. But this was due to insufficient sensitivity of the recording instruments; Davis (**2**) has recently shown that increased action potentials in the to-be-overtly-reacting arm are developed by the ready signal, and that their maxima coincides with the optimal interval between ready signal and stimulus presentation. In other words, the curve of relative antecedent tension parallels the reaction time index of optimal length of fore-period. Furthermore, Freeman and Kendall (**7**) have found that muscular tensions induced experimentally at various intervals prior to stimulus presentation influence reaction time after the manner of the Woodrow (**12**) fore-period analysis. Finally, Knott (**9**) has reported that the blocking of electroencephalograms in mental sets—which might seem to be exclusively central—is peripherally determined. These experiments, and not that of Hathaway, are the ones which are relevant to the question of the peripheral determinants of set.

In my opinion, the title of Mowrer, Rayman, and Bliss's report is a gross misnomer; and I am frankly concerned lest its

[3] Appropriate shifts in tension patterns have been shown to occur under other protracted work conditions, such as sensori-motor learning (**1**, **3**) and adaptation to 'distracting' stimuli (**5**).

uncritical reading will provide Gestalt psychologists with apparent support for their unsubstantiated doctrines of centrally maintained brain dynamics. I do not hold with these authors that "the only scientific grounds for wishing to make all psychological processes dependent upon concomitant stimulus-response sequences was the supposition that this would facilitate their empirical investigation." Nor do I see why this supposition "does not appear to have been justified." Why did Ribot propose his motor theory of attention, Feré his principle of dynamogensis, or Washburn, Guthrie and others a motor account of imagery, perception, and learning? Certainly not to initiate *new* research, but because already known facts warranted some such an explanation. A like situation holds for the explanation of set phenomena, and any advocates of exclusively central determinants have not read with sufficient thoroughness the relevant literature. Elsewhere (**6**) I have pointed out that, in general, Gestalt psychologists, in spite of their rightful insistence upon complete behavioral description and the total stimulus field, are making a serious omission by holding to central determinants alone, inasmuch as peripheral reactions, by their backlash effects, are an important and essential part of the total field. Similarly, Mowrer, Rayman, and Bliss, by their failure to consider adequately the role of differential muscular adjustments, have inadvertently threatened to throw set phenomena into the laps of Scholasticism and Rational Psychology.

(Manuscript received April 16, 1940)

REFERENCES

1. DANIEL, R. S., The distribution of muscular action potentials during maze learning, *J. exp. Psychol.*, 1939, **24**, 621–629.
2. DAVIS, R. C., Muscular tension and set, *Univ. Ind. Publ. Sci. Ser.*, 1940 (in press).
3. FREEMAN, G. L., The spread of neuro-muscular activity during mental work, *J. gen. Psychol.*, 1931, **5**, 479–494.
4. ——, The problem of set, *Amer. J. Psychol.*, 1931, **52**, 16–30.
5. ——, Changes in tension pattern and total energy expenditure during adaptation to 'distracting' stimuli, *Amer. J. Psychol.*, 1939, **52**, 354–360.
6. ——, Concerning the 'field' in 'field' psychology, *Psychol. Rev.*, 1940 (in press).
7. ——, & KENDALL, W. E., The effect upon reaction time of muscular tension induced at various preparatory intervals, *J. exp. Psychol.*, 1940 (in press).

8. HATHAWAY, S. R., An action potential study of neuromuscular relations, *J. exp. Psychol.*, 1935, **18**, 285–298.
9. KNOTT, J. R., Some effects of mental set upon the electrophysiological processes of the human cerebral cortex, *J. exp. Psychol.*, 1939, **24**, 384–405.
10. MOWRER, O. H., RAYMAN, N. N., & BLISS, E. L., Preparatory set (expectancy) . . . An experimental demonstration of its 'central' locus, *J. exp. Psychol.*, 1940, **26**, 357–372.
11. SHARP, L. H., The effects of residual tension on output and energy expenditure in muscular work. Unpublished Doctoral Dissertation, Northwestern University, June, 1940. (In preparation for press.)
12. WOODROW, H., The measurement of attention, *Psychol. Monogr.*, 1914, **17**, No. 5, 1–158.

Vol. 26, No. 6 June, 1940

Journal of Experimental Psychology

Science Library
JUN 3 1940

EDITED BY
SAMUEL W. FERNBERGER
UNIVERSITY OF PENNSYLVANIA

CONSULTING EDITORS

CHARLES W. BRAY, *Princeton University;* ELMER K. CULLER, *University of Rochester;* CLARENCE H. GRAHAM, *Brown University;* JOY P. GUILFORD, *University of Nebraska;* FRANCIS W. IRWIN, *University of Pennsylvania;* DONALD G. MARQUIS, *Yale University;* ARTHUR W. MELTON, *University of Missouri;* CARROLL C. PRATT, *Rutgers University.*

CONTENTS

Studies of Abnormal Behavior in the Rat. III. The Development of Behavior Fixations Through Frustration: N. R. F. MAIER, N. M. GLASER AND J. B. KLEE 521

Variability in Learning Under Massed and Spaced Practice: H. E. GARRETT 547

Experimental Studies in Rote-Learning Theory. VI. Comparison of Retention Following Learning to Same Criterion by Massed and Distributed Practice: C. I. HOVLAND 568

'Hypothesis' and 'Random Activity' During the Conditioning of Dogs: W. N. KELLOGG AND I. S. WOLF 588

The Relationship Between Performance Level and Bodily Activity Level: G. L. FREEMAN 602

Trial and Error Learning in Paramecium: J. W. FRENCH 609

The Variability of Extinction Scores in 'Skinner-Box' Experiments: L. G. HUMPHREYS 614

A Study of Chord Preference in a Group of Negro College Women: O. W. EAGLESON AND L. E. TAYLOR 619

Discussion: 'Central' vs. 'Peripheral' Locus of Set; A Critique of the Mowrer, Rayman, and Bliss 'Demonstration' 622

PUBLISHED MONTHLY, TWO VOLUMES PER YEAR
BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STS., LANCASTER, PA.
AND OHIO STATE UNIVERSITY, COLUMBUS, OHIO

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

PUBLICATIONS OF

THE AMERICAN PSYCHOLOGICAL ASSOCIATION

WILLARD L. VALENTINE, *Business Manager*

PSYCHOLOGICAL REVIEW

HERBERT S. LANGFELD, *Editor*
Princeton University

Contains original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.
Subscription: $5.50 (Foreign, $5.75). Single copies, $1.00.

PSYCHOLOGICAL BULLETIN

JOHN A. McGEOCH, *Editor*
State University of Iowa

Contains critical reviews of books and articles, psychological news and notes, university notices, and announcements. Appears monthly (10 issues), the annual volume comprising about 665 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.
Subscription: $7.00 (Foreign, $7.25). Single copies, 75c.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

S. W. FERNBERGER, *Editor*
University of Pennsylvania

Contains original contributions of an experimental character. Appears monthly (since January, 1937), two volumes per year, each volume of six numbers containing about 625 pages.
Subscription: $14.00 ($7.00 per volume; Foreign, $7.25). Single copies, $1.25.

PSYCHOLOGICAL ABSTRACTS

WALTER S. HUNTER, *Editor*
Brown University

Appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.
Subscription: $7.00 (Foreign, $7.25). Single copies, 75c.

PSYCHOLOGICAL MONOGRAPHS

JOHN F. DASHIELL, *Editor*
University of North Carolina

Consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.
Subscription: $6.00 per volume (Foreign, $6.30).

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

GORDON W. ALLPORT, *Editor*
Harvard University

Appears quarterly, January, April, July, October, the four numbers comprising a volume of 560 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.
Subscription: $5.00 (Foreign, $5.25). Single copies, $1.50.

COMBINATION RATES

Review and Bulletin: $11.00 (Foreign, $11.50).
Review and J. Exp. (2 vols.): $17.00 (Foreign, $17.75).
Bulletin and J. Exp. (2 vols.): $18.50 (Foreign, $19.25).
Review, Bulletin, and J. Exp. (2 vols.): $23.00 (Foreign, $24.00).

Subscriptions, orders, and business communications should be sent to

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

THE OHIO STATE UNIVERSITY, COLUMBUS, OHIO